UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Computational Biology and Complex Systems Research Group

CMCiB Comparative Medicine & Bioimage Centre of Catalonia

IGTP Germans Trias i Pujol Research Institute

Daily report 29-05-2020

Analysis and prediction of COVID-19 for EU-EFTA-UK and other countries

Situation report 76

Contact: clara.prats@upc.edu

With the financial support of: MINISTERIO DE CIENCIA E INNOVACIÓN "la Caixa"

Foreword

The present report aims to provide a comprehensive picture of the **pandemic situation of COVID-19** in the EU countries, and to be able to foresee the situation in the next coming days.

We employ an **empirical model**, verified with the evolution of the number of confirmed cases in previous countries where the epidemic is close to conclude, including all provinces of China. The model does not pretend to interpret the causes of the evolution of the cases but to permit the **evaluation of the quality of control measures made in each state** and a **short-term prediction of trends**. Note, however, that the effects of the measures' control that start on a given day are not observed until approximately 7-10 days later.

The model and predictions are based on two parameters that are daily fitted to available data:

- ✓ *a*: the velocity at which spreading specific rate slows down; the higher the value, the better the control.
- ✓ *K*: the final number of expected cumulated cases, which cannot be evaluated at the initial stages because growth is still exponential.

We show an individual report with 8 graphs and a table with the **short-term predictions** for different countries and regions. We are adjusting the model to **countries and regions** with at least 4 days with more than 100 confirmed cases and a current load over 200 cases. The **predicted period** of a country depends on the number of datapoints over this 100 cases threshold, and is of 5 days for those that have reported more than 100 cumulated cases for 10 consecutive days or more. For short-term predictions, we assign higher weight to last 3 points in the fittings, so that changes are rapidly captured by the model. The whole methodology employed in the inform is explained in the last pages of this document.

In addition to the individual reports, the reader will find an initial dashboard with a brief analysis of the situation in EU-EFTA-UK countries, some summary figures and tables as well as **long-term predictions** for some of them, when possible. These long-term predictions are evaluated without different weights to data-points. We also discuss a specific issue every day.

Martí Català
Pere-Joan Cardona, PhD
Comparative Medicine and Bioimage Centre of Catalonia; Institute for Health Science Research Germans Trias i Pujol

Clara Prats, PhD
Sergio Alonso, PhD
Enric Álvarez, PhD
Miquel Marchena
David Conesa
Daniel López, PhD
Computational Biology and Complex Systems; Universitat Politècnica de Catalunya - BarcelonaTech

With the collaboration of: Guillem Álvarez, Oriol Bertomeu, Laura Dot, Lavínia Hriscu, Helena Kirchner, Daniel Molinuevo, Pablo Palacios, Sergi Pradas, David Rovira, Xavier Simó, Tomás Urdiales

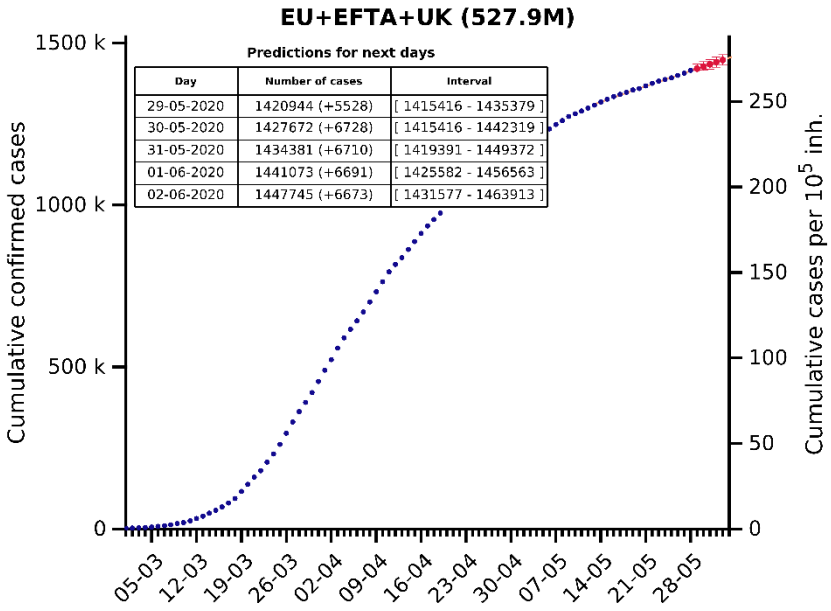
PJC and MC received funding from "la Caixa" Foundation (ID 100010434), under agreement LCF/PR/GN17/50300003; CP, DL, SA, MC, received funding from Ministerio de Ciencia, Innovación y Universidades and FEDER, with the project PGC2018-095456-B-I00;

Disclaimer: These reports have been written by declared authors, who fully assume their content. They are submitted daily to the European Commission, but this body does not necessarily share their analyses, discussions and conclusions.

(0) Executive summary – Dashboard

Global EU+EFTA+UK trends and needs

Spain, Italy and France, the three largest countries in southwestern Europe, have behaved quite similarly. In this brief analysis we are going to look at them as a whole. It is an interesting exercise because it provides an image that can be compared with countries like the USA or Brazil.



The three countries together account for about 173 million inhabitants. To date, 618,709 cases have been reported. Its temporal behavior fits perfectly with the Gompertz model, with an R2 of 0.999 and an expected K of 624,215 cases. Today, we have exceeded the threshold of 99% of the forcasted total cases. This group of countries had the peak in new cases between March 26 and April 2, when they reported more than 18,000 new cases a day, and currently have approximately 1,500 new cases daily (0.87 per 100,000 inhabitants). It is to be hoped that this value may continue to fall, although it would not be unlikely that it would stabilize in short.

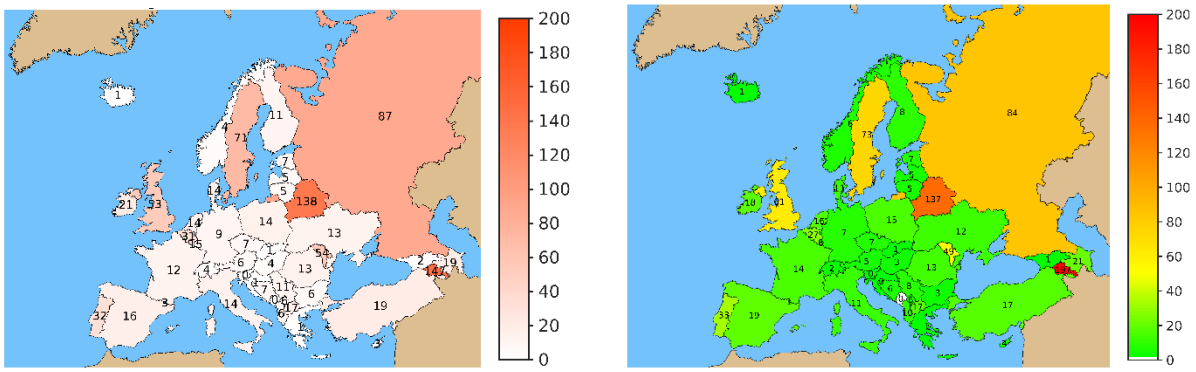
We have not considered the latest data provided by France which seems to need some explanation, as the value is much higher than expected.

The analysis is focused on discussing fluctuations in European countries datasets..

Trends for specific countries

The situation and trends of countries is similar to the one reported yesterday, except for a sudden increase in last data point of **France**. Both **UK** and **France** have experimented a one-day increase this week. Therefore, ρ_7 will be affected for a few days. **Sweden** remains reporting 500-600 daily new cases, with a ρ_7 of 1.04. **Netherlands** and Portugal seem to be stabilizing at the level of 200 cases daily ($\rho_7 \approx 1$).

The map in the left shows current $\mathbf{A_{14}}$. The map in the right shows current **EPG**.



Situation and trends per country

Table of current situation in EU countries. Colour scale is relative except when indicated, this means that it is applied independently to each column, and distinguishes best (green) form worst (red) situations according to each of the variables. Last column (**EPG_{EST}**) indicates EPG assessed with **estimated real 14-day attack rate** (see report from 22/04 for details). **EPG_{REP}** is calculated with **data reported by countries**. EPG_{REP} and EPG_{EST} **cannot be compared between them** because scales are different, but can be independently used for estimating risk of countries according to reported or estimated real situation, respectively.

	Reported data								Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	Estimated active cases (last 14 days)	Estimated 14-day attack rate /10^5 inh.	ρ_7 [1]	EPG_{REP} [2]	EPG_{EST} [3]
United Kingdom	269.127	405,1	37.837	56,9	35.976	54,1	523.939	771,8	1,15	62	889
Spain	237.906	513,3	27.119	58,5	8.366	18,1	97.334	207,0	1,04	19	216
Italy	231.732	389,9	33.142	55,8	8.636	14,5	125.381	207,4	0,80	12	165
Germany	180.458	220,3	8.450	10,3	7.306	8,9	36.223	43,2	0,82	7	36
France	149.071	230,3	28.662	44,3	7.715	11,9	153.943	235,8	1,21	14	286
Belgium	57.849	509,3	9.388	82,7	3.561	31,4	58.710	506,6	0,88	28	444
Netherlands	45.950	270,5	5.903	34,8	2.469	14,5	32.509	189,7	1,11	16	211
Sweden	35.727	363,2	4.266	43,4	7.145	72,6	92.352	914,4	1,04	75	947
Portugal	31.596	304,6	1.369	13,2	3.277	31,6	14.763	144,8	1,01	32	146
Switzerland	30.713	358,4	1.654	19,3	333	3,9	1.792	20,7	0,63	2	13
Ireland	24.841	525,6	1.639	34,7	1.014	21,5	7.395	149,8	0,88	19	131
Poland	22.825	59,7	1.038	2,7	5.210	13,6	26.875	71,0	1,09	15	77
Romania	18.791	95,0	1.229	6,2	2.544	12,9	17.723	92,1	0,98	13	90
Austria	16.543	189,9	668	7,7	538	6,2	2.205	24,5	0,80	5	19
Denmark	11.512	201,5	568	9,9	799	14,0	3.960	68,4	0,81	11	55
Czech Republic	9.140	86,1	319	3,0	789	7,4	2.988	27,9	0,95	7	26
Norway	8.401	156,5	236	4,4	226	4,2	648	12,0	1,42	6	17
Finland	6.743	122,5	313	5,7	598	10,9	3.018	54,5	0,77	8	42
Luxembourg	4.008	695,8	110	19,1	93	16,1	262	41,9	0,54	9	23
Hungary	3.841	39,4	517	5,3	424	4,3	5.955	61,6	0,75	3	46
Greece	2.906	26,0	175	1,6	136	1,2	787	7,5	0,84	1	6
Bulgaria	2.485	34,8	136	1,9	385	5,4	2.187	31,5	0,55	3	17
Croatia	2.245	53,3	102	2,4	24	0,6	120	2,9	0,70	0	2
Estonia	1.851	141,1	66	5,0	93	7,1	NA	NA	1,05	7	NA
Iceland	1.805	495,5	10	2,7	3	0,8	NA	NA	0,86	1	NA
Lithuania	1.656	56,9	68	2,3	145	5,0	NA	NA	0,93	5	NA
Slovakia	1.520	27,9	28	0,5	43	0,8	NA	NA	1,67	1	NA
Slovenia	1.473	70,9	108	5,2	8	0,4	61	2,9	1,14	0	3
Latvia	1.061	53,8	24	1,2	99	5,0	NA	NA	0,91	5	NA
Cyprus	941	80,4	17	1,5	34	2,9	NA	NA	1,38	4	NA
Malta	616	143,6	7	1,6	94	21,9	NA	NA	NA	NA	NA
Liechtenstein	83	215,3	1	2,6	0	0,0	NA	NA	NA	NA	NA
	Scale										
	Worst	Worst	Worst	Worst	Worst	Worst	Worst	Worst	2,0	100	1000
	Best	Best	Best	Best	Best	Best	Best	Best	0,0	0	0

[1] ρ_7 is the average of 7 consecutive ρ, but can still fluctuate. [2,3] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

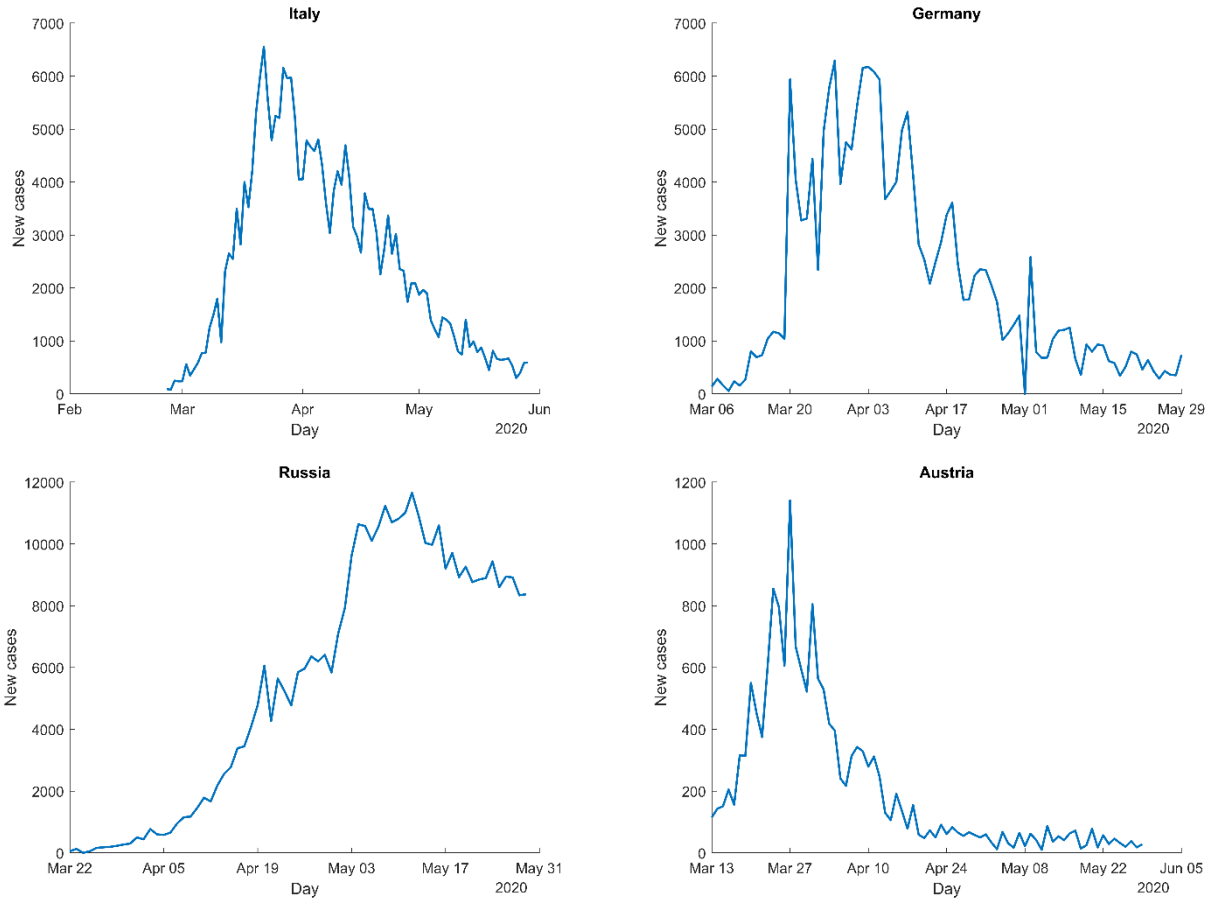
Highlights for countries with highest number of reported cases

- ✓ Spain is reviewing all historical data and reports a decrease in deaths as a consequence of this revision. There are also a few inconsistences in reported cases that will persist until complete revision is finished.
- ✓ France shows a one-day increase. Last sudden rise was due to the incorporation of a new laboratory, the explanation could be similar this time (to be confirmed).
- ✓ UK and France's ρ_7 are affected by this week's sudden increases. Nevertheless, they would be following previous trends.

Analysis: Study of the fluctuations in the reporting of new cases in the EU countries.

A new outbreak in the epidemic of Covid19 may begin with a small increase in the number of cases. As we have previously shown in the report of Monday[1] (Report #72), the **reporting of new cases every day is superposed to a large amount of different types of noises and disturbances, like the weekend effect or others, that could mask early signals of a secondary outbreak**. We have extended our previous study on the relative fluctuations in the reporting of new cases to a large number of countries trying to extract a quantitative view of the situation in different countries by simple comparison.

We have employed the method previously described to the set of data for the countries in the EU+EFTA+UK. For the visualization of the fluctuations we show the **new cases reported in Italy, Germany, Russia and Austri**a. We can observe that, while some of the countries (Italy and Russia) are more coherent with the everyday reporting data and not much deviation is observed, the other two countries (Germany and Austria) report the data with more fluctuations.



Next, we define the ***fluctuations*** for each country as the **difference between the number of new cases today and yesterday**:

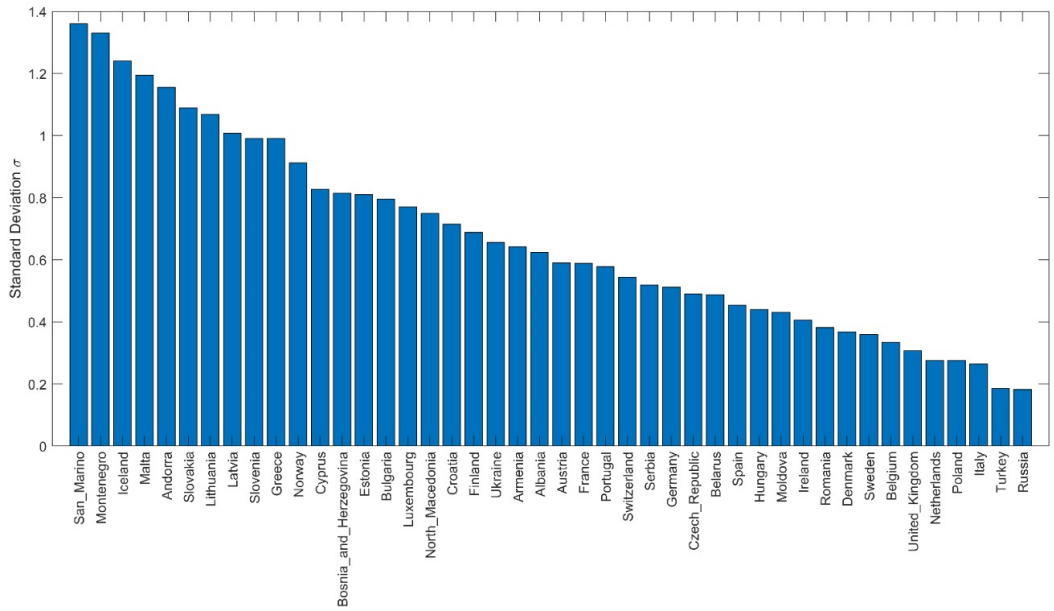
$$fluctuations = New(t) - New(t-1).$$

Let us define the ***relative fluctuations*** (RF) as the fluctuations divided by the mean number of new cases today and yesterday:

[1] https://upcommons.upc.edu/handle/2117/188961

$$RF(t) = \frac{fluctuations}{\frac{1}{2}\big(New(t) + New(t-1)\big)} = \frac{New(t) - New(t-1)}{\frac{1}{2}\big(New(t) + New(t-1)\big)}$$

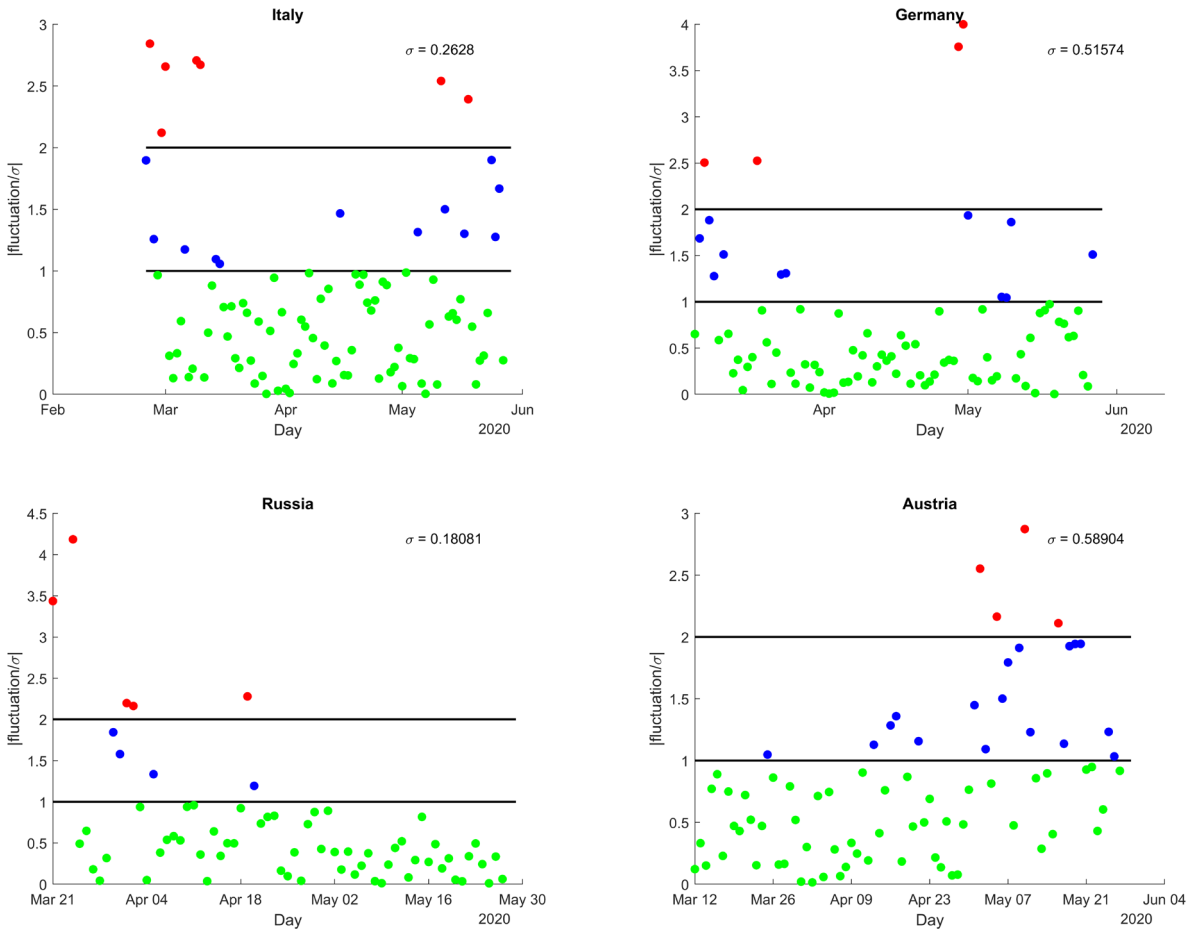
As seen on Monday (Report #72), we can compute the variance of the relative fluctuations (RF). In the next figure **we show the resulting variance for all the countries** under consideration. The higher the bar, the higher the relative fluctuations.



The value of σ defines the characteristic relative fluctuations of the country[2]. The countries in the figure showing large values of σ correspond to countries with small populations, where, furthermore, the epidemic is at the last stage and therefore the values of $New(t)$ are small making the RF appear very large because its effect in the denominator.

Relative fluctuations divided by sigma (remember that each country has a different value for sigma) follow a complex distribution, see the four examples in the next figure. It shows the relative fluctuations divided by sigma for each day of the epidemic. Three types of fluctuations can be distinguished: small fluctuations (in green) alternate with moderate fluctuations (in blue) and with some larger disturbances (in red).

[2] If the fluctuations would follow a Gaussian distribution, a value of σ=1 would mean that the 32% of the values of the distribution of RF correspond to variations $New(t) - New(t-1) > New(t-1)$ and the 68% of the values correspond to $New(t) - New(t-1) < New(t-1)$. Equivalently, for an arbitrary value of σ we have that around 32% of the values correspond to $New(t) - New(t-1) > \sigma \cdot New(t-1)$ and the 68% of the values correspond to $New(t) - New(t-1) < \sigma \cdot New(t-1)$.



It is relatively common in the series of the new cases in the countries that, suddenly, a big update appears with some hundreds or thousands of forgotten cases (see France's reported cases today). Such fluctuations are typically artificial and, in order to avoid them, we have released the points outside 5 times σ to avoid a big bias of the country for a particular day.

The calculation of the parameter σ for each country therefore permits the estimation of the characteristic fluctuations on the reporting protocol and therefore **σ defines a threshold for an abnormal value of new cases to be consider as a candidate of new outbreak**. In order to define a threshold, we consider twice the value of the value of σ, then **if the fluctuation produced by a new value New* is bigger than 2σ, we consider as candidate of new outbreak**. Obviously next days are critical to decide if the perturbation grows or decline to the anterior dynamics. It means that given a value for cases reported today (New), the cases reported tomorrow will produce a fluctuation

$$RF(t) = \frac{New * -New(t)}{New(t)}$$

which may be dangerous if RF> 2σ and therefore the limit for the reported cases of tomorrow is

$$New_{th} = 2\sigma New(t) + New(t)$$

We have calculated the values of the difference of cases $(2\sigma New(t))$ **and the new value of cases** $(2\sigma New(t) + New(t))$ **which may be worrying for the given actual values of the daily new cases in the different countries in EU+EFTA+UK**. In the next table, we show the values of these quantities for the different countries. Obviously, the countries with very small values of new cases, a small difference on the new cases reported may substantially increase the relative fluctuation. In the countries at the final state of the epidemic,

it may be useful to perform an average of the new cases on the last 7 or 10 days, which is not possible in the beginning of the epidemic because the large daily increase of cases.

Country	New cases (N)	Standard deviation (σ)	Worrying increase (2σN)	Number of worrying cases (N+2σN)
San Marino	2	1.36	5	7
Montenegro	0	1.33	0	0
Iceland	1	1.26	3	4
Malta	3	1.19	7	10
Andorra	0	1.16	0	0
Slovakia	4	1.09	9	13
Lithuania	9	1.06	19	28
Latvia	4	1.00	8	12
Greece	7	0.99	14	21
Slovenia	2	0.98	4	6
Norway	19	0.91	34	53
Cyprus	1	0.86	2	3
Bulgaria	13	0.83	22	35
Bosnia and Herz.	23	0.81	37	60
Estonia	9	0.81	15	24
Luxembourg	7	0.76	11	18
Croatia	1	0.75	2	3
North Macedonia	32	0.75	48	80
Ukraine	399	0.70	560	959
Finland	58	0.68	79	137
Armenia	407	0.64	518	925
Albania	24	0.62	30	54
France	1758	0.62	2173	3931
Austria	23	0.59	27	50
Switzerland	18	0.59	21	39
Portugal	295	0.57	339	634
Serbia	37	0.52	38	75
Germany	547	0.52	564	1111
Czech Republic	45	0.49	44	89
Belarus	900	0.48	870	1770
Spain	569	0.46	524	1093
Hungary	24	0.44	21	45
Moldova	210	0.43	180	390
Ireland	53	0.41	43	96
Romania	181	0.38	138	319
Denmark	42	0.37	31	73
Sweden	644	0.36	459	1103
Belgium	197	0.34	134	331
United Kingdom	1950	0.31	1190	3140
Poland	376	0.27	206	582
Netherlands	186	0.27	102	288
Italy	589	0.26	310	899
Turkey	1109	0.18	410	1519
Russia	8355	0.18	3021	11376

Situation and trends in other countries

Table of current situation in a sample of non-EU countries. Colour scale is relative except when indicated, this means that it is applied independently to each column, and distinguishes best (green) form worst (red) situations according to each of the variables. EPG_{REP} and EPG_{EST} **cannot be compared between them** because scales are different, but can be independently used for estimating risk of countries according to reported or estimated real situation, respectively.

	Reported data								Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	Estimated active cases (last 14 days)	Estimated 14-day attack rate /10^5 inh.	ρ_7 [1]	EPG_{REP} [2]	EPG_{EST} [3]
United States of America	1.721.750	520,2	101.617	30,7	303.861	91,8	1.930.295	583,2	0,95	87	553
Brazil	438.238	206,2	26.754	12,6	235.320	110,7	1.733.710	815,6	1,16	128	943
Russia	379.051	259,7	4.142	2,8	126.806	86,9	NA	NA	0,97	84	NA
India	165.799	12,3	4.706	0,3	83.829	6,2	284.054	21,0	1,20	7	25
Iran	143.849	171,3	7.627	9,1	29.316	34,9	164.482	195,8	0,99	35	194
Peru	141.779	430,0	4.099	12,4	61.175	185,5	209.293	634,8	1,15	214	732
Canada	88.501	234,5	6.877	18,2	15.100	40,0	134.647	356,8	0,94	38	336
Chile	86.943	454,8	890	4,7	49.903	261,1	107.173	560,6	1,36	356	764
Mexico	81.400	63,1	9.044	7,0	38.805	30,1	560.111	434,4	1,20	36	522
Saudi Arabia	80.185	230,3	441	1,3	33.316	95,7	NA	NA	0,87	84	NA
Pakistan	64.028	29,0	1.317	0,6	26.810	12,1	62.778	28,4	0,98	12	28
Qatar	50.914	1.767,2	33	1,1	22.642	785,9	NA	NA	1,09	860	NA
Belarus	39.858	421,8	219	2,3	13.086	138,5	NA	NA	0,99	137	NA
Ecuador	38.471	218,1	3.313	18,8	7.969	45,2	77.523	439,4	0,78	35	344
Argentina	14.689	32,5	508	1,1	7.568	16,7	45.799	101,3	1,68	28	170
	Scale										
	Worst	Worst	Worst	Worst	Worst	Worst	Worst	Worst	2,0	100	1000
	Best	Best	Best	Best	Best	Best	Best	Best	0,0	0	0

[1] ρ_7 is the average of 7 consecutive ρ, but can still fluctuate. [2,3] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

Disclaimer: estimated active cases and estimated 14-day attack rate are assessed by assuming a lethality of 1 % (see report from 20 to 24 April, #37-41). This value can change in countries where suspicious deaths are reported as well (real values would be lower) and in countries where incidence among elderly people was minor (real values would be higher).

Time indicators by country

These tables summarize a few time indicators for each country: time since 50 cases were reported, time interval between an attack rate of $1/10^5$ inhabitants and an attack rate of $10/10^5$ inhabitants, and time interval between attack rates of 10 to 100 per 10^5 inhabitants (only for countries that have overtaken this threshold).

EU+EFTA+UK countries

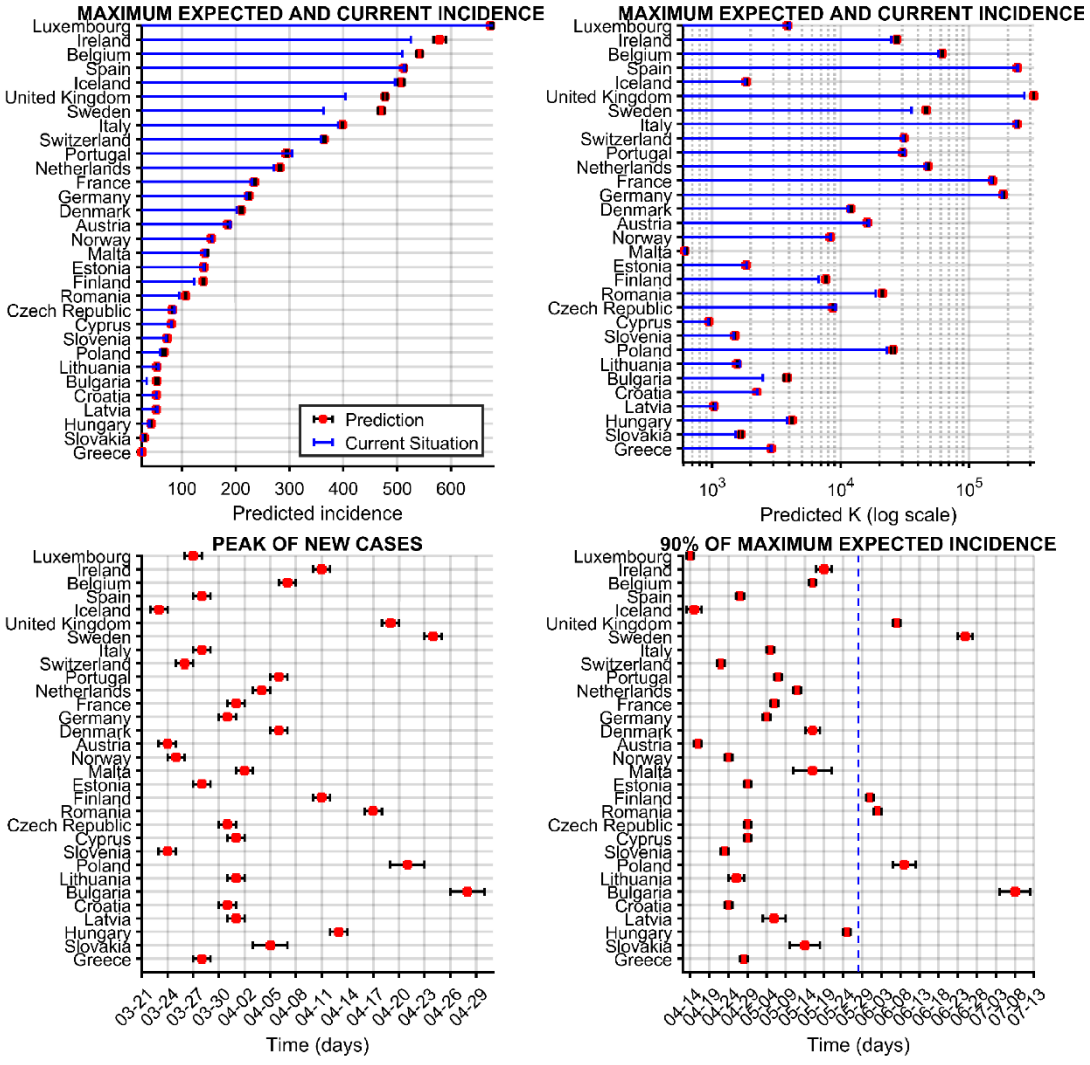
Countries	Days since the first 100 cases	Time interval between 1 and 10 cases / 10^5 inh. (days)	Time interval between 10 and 100 cases / 10^5 inh. (days)
Italy	96	11	16
Germany	90	12	17
France	89	10	20
Spain	89	8	12
United Kingdom	85	10	12
Belgium	84	11	14
Netherlands	84	11	20
Sweden	84	10	28
Norway	84	2	7
Switzerland	84	8	11
Austria	82	10	14
Denmark	81	4	30
Czech Republic	78	11	NA
Finland	78	12	46
Greece	78	18	NA
Iceland	78	5	15
Portugal	77	9	15
Slovenia	77	6	NA
Estonia	76	5	30
Ireland	76	8	18
Poland	76	17	NA
Romania	76	15	NA
Luxembourg	73	6	7
Slovakia	72	24	NA
Bulgaria	71	30	NA
Croatia	71	12	NA
Hungary	70	20	NA
Latvia	70	12	NA
Lithuania	69	9	NA
Malta	68	9	35
Cyprus	67	12	NA

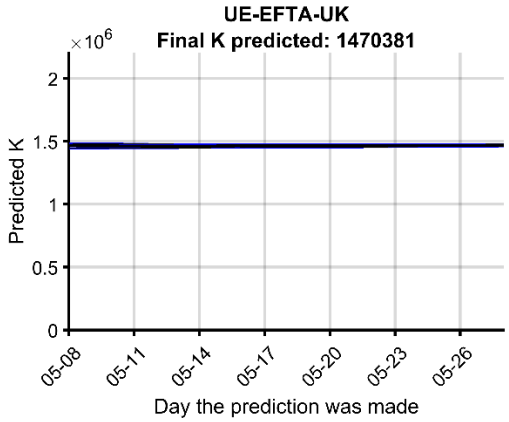
Other countries

Countries	Days since the first 100 cases	Time interval between 1 and 10 cases / 10^5 inh. (days)	Time interval between 10 and 100 cases / 10^5 inh. (days)
Iran	93	11	42
United States of America	88	8	15
Canada	79	11	27
Qatar	79	3	31
Brazil	76	20	34
Saudi Arabia	75	21	29
Chile	74	13	36
Pakistan	74	35	NA
India	74	38	NA
Russia	73	15	24
Peru	73	18	22
Ecuador	73	10	30
Mexico	72	25	NA
Argentina	71	39	NA
Belarus	60	10	18

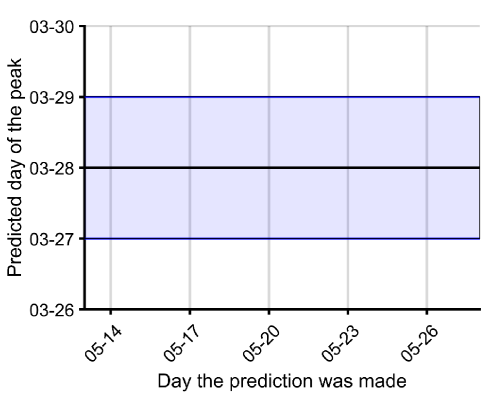
Long-term predictions

Evaluated with the **whole historical series**. See figure in the next page. Up-left: Predictions of maximum incidences per country (total final expected attack rate per 10^5 inh.). Up-right: Predictions of maximum absolute number of cases per country (K, in log scale). Blue lines indicate current situation. Bottom-left: Time in which peak in new cases was achieved / will be achieved. Bottom-right: Time at which 90 % of K was achieved / will be achieved. Blue dotted line indicates current date. At the end, predicted K for whole EU+EFTA+UK.



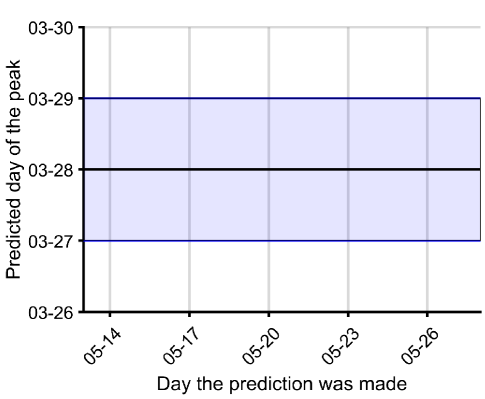


2020-05-28


03-30
03-29
03-28
03-27
03-26
Predicted day of the peak
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

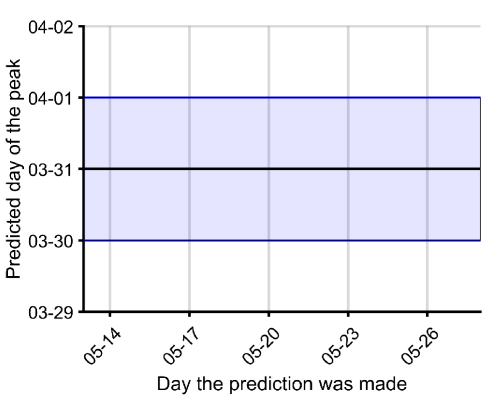
Spain
04-29
04-28
04-27
04-26
04-25
04-24
Predicted day for 90% of K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

×10^5
3.5
3
2.5
2
1.5
1
0.5
0
Predicted K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made


03-30
03-29
03-28
03-27
03-26
Predicted day of the peak
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

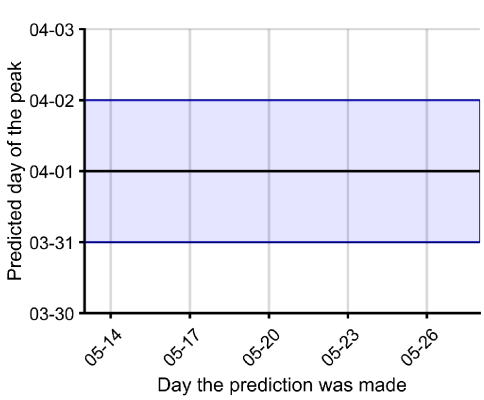
Italy
05-07
05-06
05-05
05-04
05-03
Predicted day for 90% of K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

×10^5
3.5
3
2.5
2
1.5
1
0.5
0
Predicted K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made


04-02
04-01
03-31
03-30
03-29
Predicted day of the peak
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

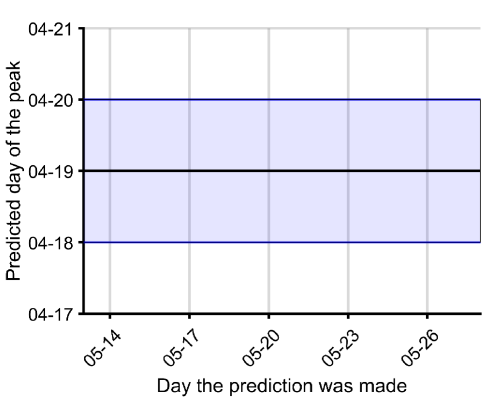
Germany
05-07
05-06
05-05
05-04
05-03
05-02
05-01
Predicted day for 90% of K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

×10^5
2.5
2
1.5
1
0.5
0
Predicted K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made


04-03
04-02
04-01
03-31
03-30
Predicted day of the peak
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

France
05-10
05-09
05-08
05-07
05-06
05-05
05-04
Predicted day for 90% of K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

×10^5
2
1.5
1
0.5
0
Predicted K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made


04-21
04-20
04-19
04-18
04-17
Predicted day of the peak
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

United Kingdom
06-09
06-08
06-07
06-06
06-05
Predicted day for 90% of K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

×10^5
4
3
2
1
0
Predicted K
05-14
05-17
05-20
05-23
05-26
Day the prediction was made

Situation and trends in Italian regions[3]

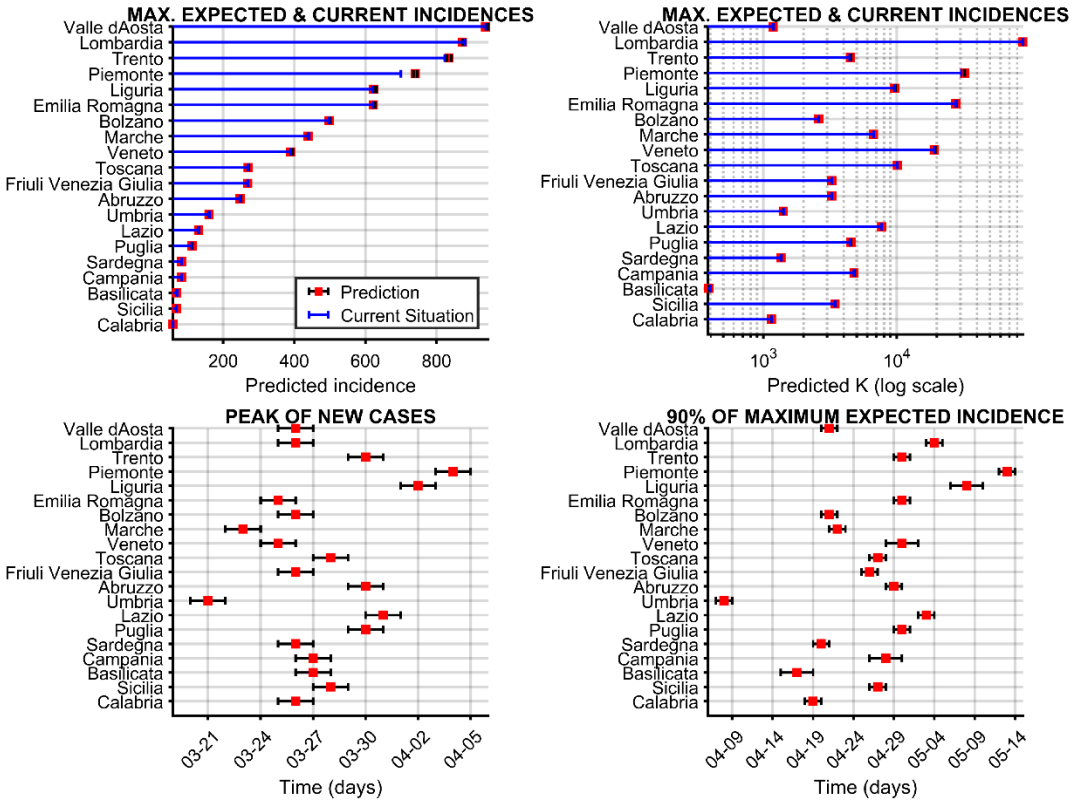
Situation and trends

	Reported data								Indexes		
Country	Cumulative cases	Attack rate /10^5 inh.	Cumulative deaths	Mortality /10^5 inh.	Active cases (last 14 days)	14-day attack rate /10^5 inh.	Estimated active cases (last 14 days)	Estimated 14-day attack rate /10^5 inh.	ρ_7 [(1)]	EPG_{REP} [(2)]	EPG_{EST} [(3)]
Lombardia	88,537	881.7	16,012	159.5	4,418	44.0	80,886	804.0	0.92	41	742
Piemonte	30,501	700.1	3,851	88.4	1,155	26.5	14,809	340.0	0.72	19	246
Emilia Romagna	27,739	622.0	4,102	92.0	629	14.1	9,318	209.0	0.81	11	169
Veneto	19,134	390.0	1,906	38.9	245	5.0	2,500	50.9	0.56	3	29
Toscana	10,088	270.5	1,031	27.6	205	5.5	2,101	56.3	0.62	3	35
Liguria	9,619	620.3	1,452	93.6	559	36.0	8,621	555.8	0.82	30	455
Lazio	7,709	131.1	721	12.3	345	5.9	3,443	58.6	0.64	4	38
Marche	6,723	440.8	986	64.6	104	6.8	1,487	97.5	0.73	5	71
Campania	4,787	82.5	411	7.1	133	2.3	1,146	19.8	0.77	2	15
Puglia	4,482	111.2	500	12.4	116	2.9	1,315	32.6	0.99	3	32
Trento	4,428	413.0	462	43.1	110	10.3	1,130	210.1	0.68	7	143
Sicilia	3,440	68.8	272	5.4	66	1.3	529	10.6	0.70	1	7
Friuli Venezia Giulia	3,267	268.8	333	27.4	92	7.6	964	79.3	0.91	7	72
Abruzzo	3,237	246.8	404	30.8	89	6.8	1,056	80.5	0.49	3	40
Bolzano	2,595	2,415.4	291	270.9	17	15.8	191	36.7	0.50	8	18
Umbria	1,431	162.2	76	8.6	9	1.0	NA	NA	0.44	0	NA
Sardegna	1,356	82.7	130	7.9	8	0.5	73	4.5	0.15	0	1
Valle dAosta	1,182	941.0	143	113.8	10	8.0	141	112.0	1.21	10	136
Calabria	1,158	59.5	97	5.0	14	0.7	NA	NA	0.41	0	NA
Molise	436	142.7	22	7.2	29	9.5	NA	NA	0.60	6	NA
Basilicata	399	70.9	27	4.8	10	1.8	NA	NA	1.26	2	NA
	Scale										
	Worst	Worst	Worst	Worst	Worst	Worst	Worst	Worst	2.0	100	1000
	Best	Best	Best	Best	Best	Best	Best	Best	0.0	0	0

[(1)] ρ_7 is the average of 7 consecutive ρ, but can still fluctuate. [(2,3)] EPG stands for Effective Growth Potential. EPG_{REP} is obtained by multiplying attack rate of last 14 days per 10^5 inhabitants (i.e. density of cases) by ρ_7 (a value related with effective reproduction number and that, therefore, determines the dynamics for subsequent days). EPG_{EST} is obtained by multiplying estimated real attack rate of last 14 days per 10^5 inhabitants by ρ_7.

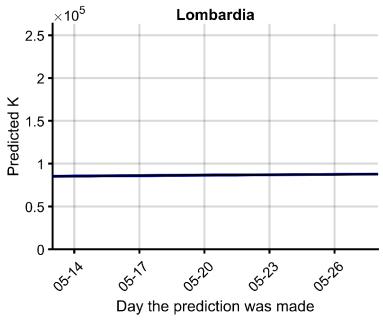
Disclaimer: estimated active cases and estimated 14-day attack rate are assessed by assuming a lethality of 1 % (see report from 20 to 24 April, #37-41). This value can change in countries where suspicious deaths are reported as well (real values would be lower) and in countries where incidence among elderly people was minor (real values would be higher).

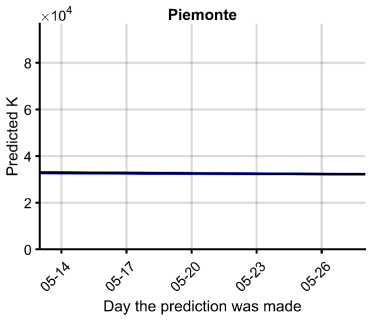
Long-term predictions

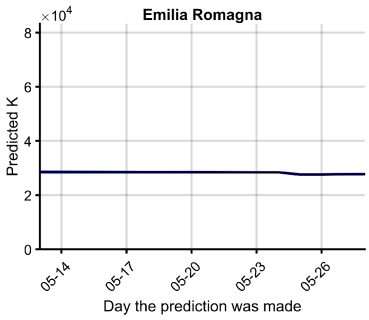


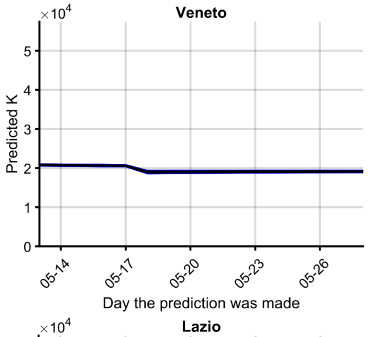
[3] **Spain:** Historical series have not been updated. Therefore, regional analysis is not shown

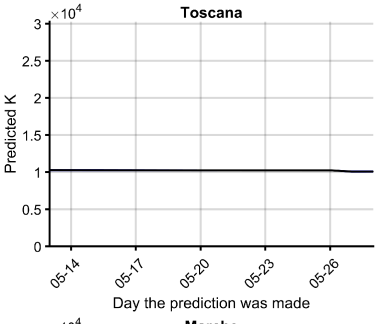
2020-05-29

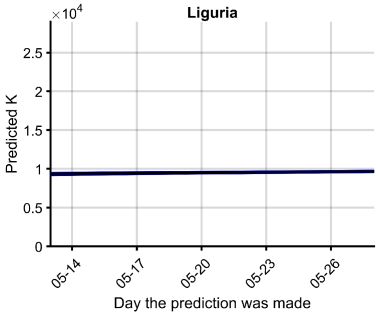

Lombardia
×10^5
Predicted K
Day the prediction was made

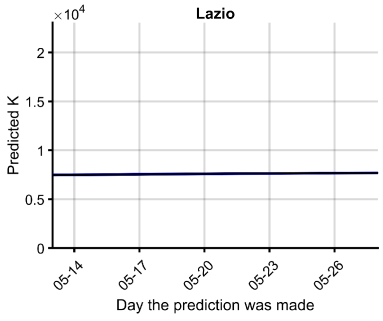

Piemonte
×10^4
Predicted K
Day the prediction was made

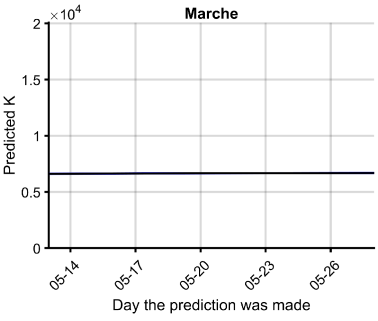

Emilia Romagna
×10^4
Predicted K
Day the prediction was made

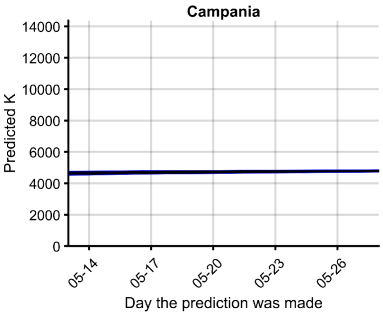

Veneto
×10^4
Predicted K
Day the prediction was made

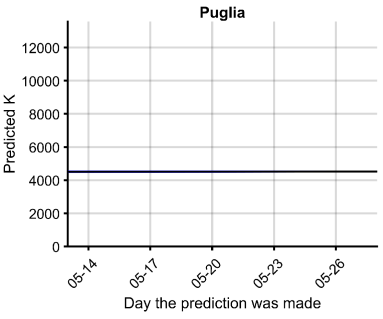

Toscana
×10^4
Predicted K
Day the prediction was made

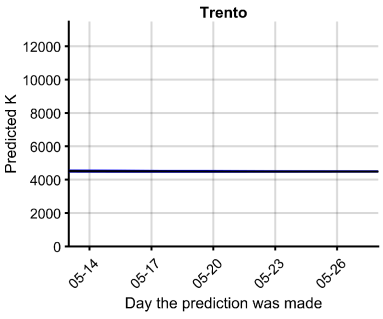

Liguria
×10^4
Predicted K
Day the prediction was made

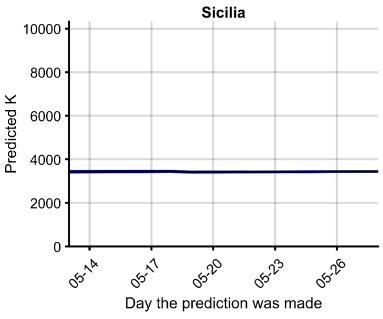

Lazio
×10^4
Predicted K
Day the prediction was made

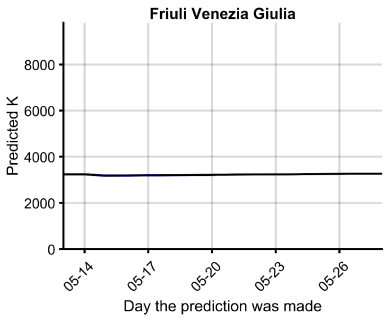

Marche
×10^4
Predicted K
Day the prediction was made

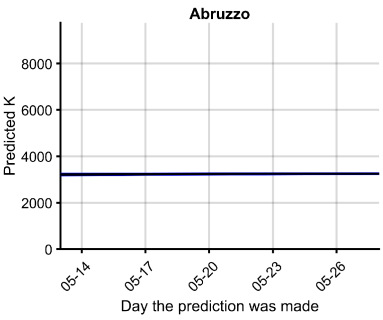

Campania
Predicted K
Day the prediction was made

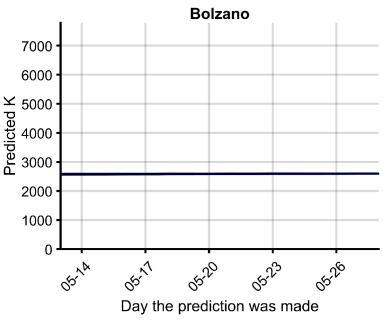

Puglia
Predicted K
Day the prediction was made

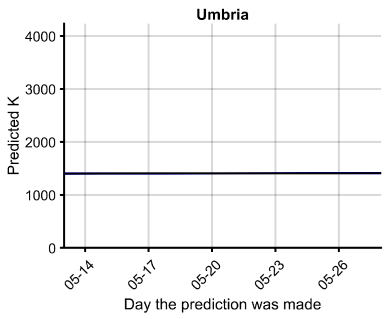

Trento
Predicted K
Day the prediction was made

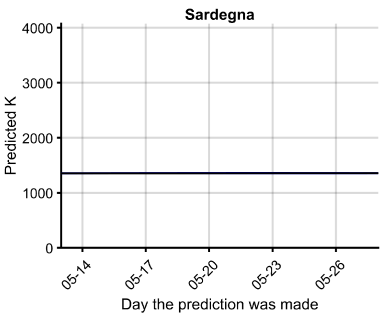

Sicilia
Predicted K
Day the prediction was made

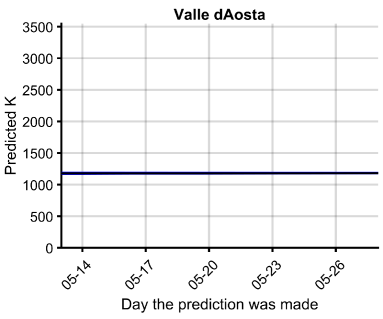

Friuli Venezia Giulia
Predicted K
Day the prediction was made

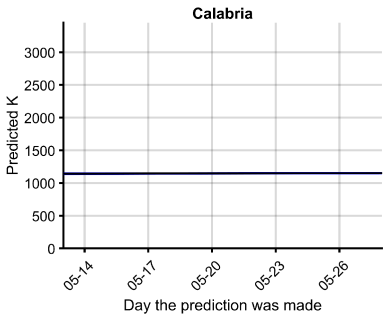

Abruzzo
Predicted K
Day the prediction was made

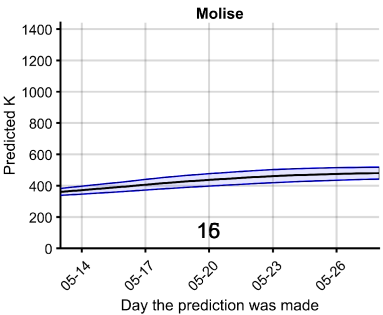

Bolzano
Predicted K
Day the prediction was made

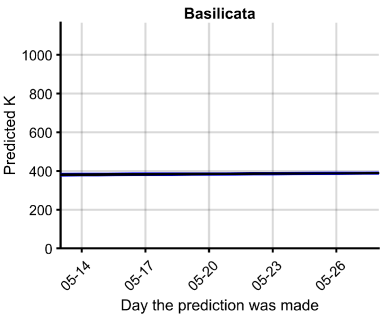

Umbria
Predicted K
Day the prediction was made


Sardegna
Predicted K
Day the prediction was made


Valle dAosta
Predicted K
Day the prediction was made


Calabria
Predicted K
Day the prediction was made

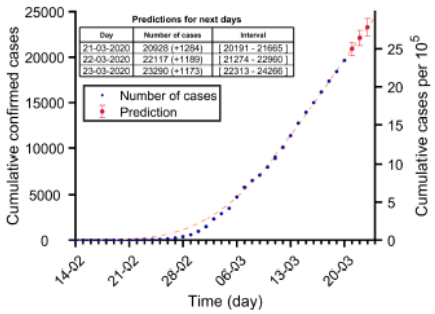

Molise
Predicted K
Day the prediction was made

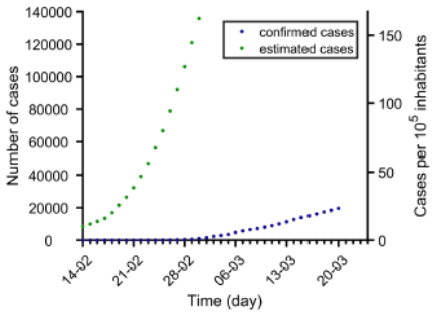

Basilicata
Predicted K
Day the prediction was made

Legend: Countries' reports details

Iran 20-03-2020. Population: 83.7M. Current cumulated incidence: $23/10^5$

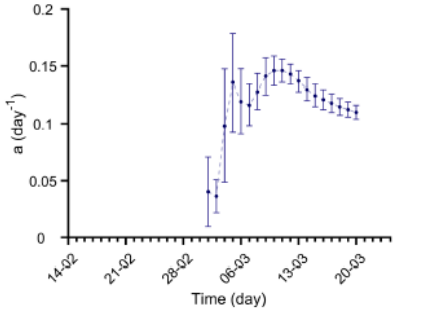
Confirmed cases: data (blue), model fitted (dashed line), predictions (red points and table)

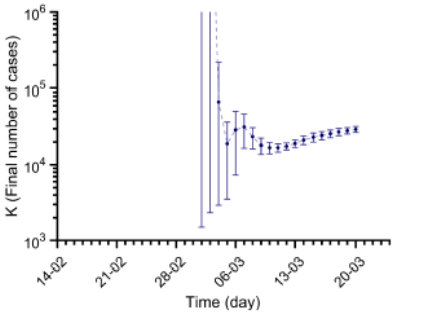




Estimated cases using death rate (see Methods)

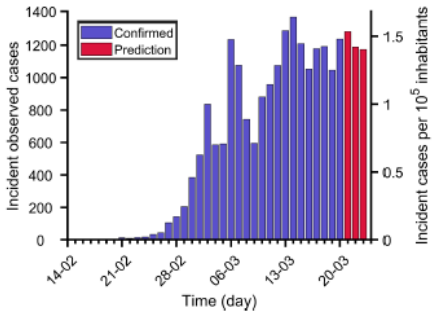
Fitted *a* value using points prior to each date

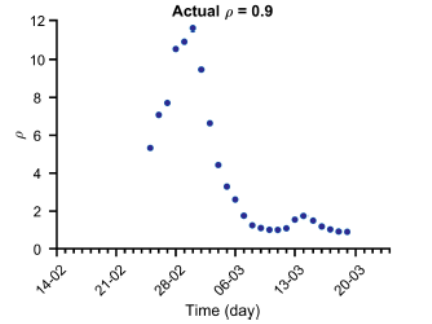




Fitted *K* value using points prior to each date

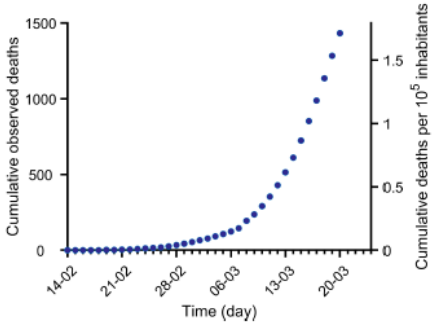
Reported and predicted new cases

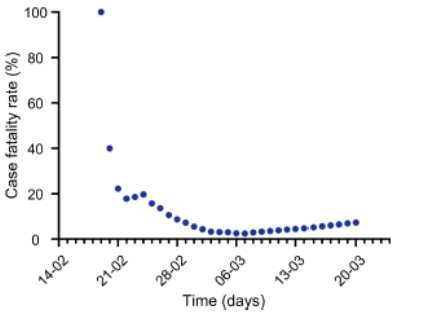




Evolution of ρ, a parameter related with Reproduction number (see Methods)

Reported deaths



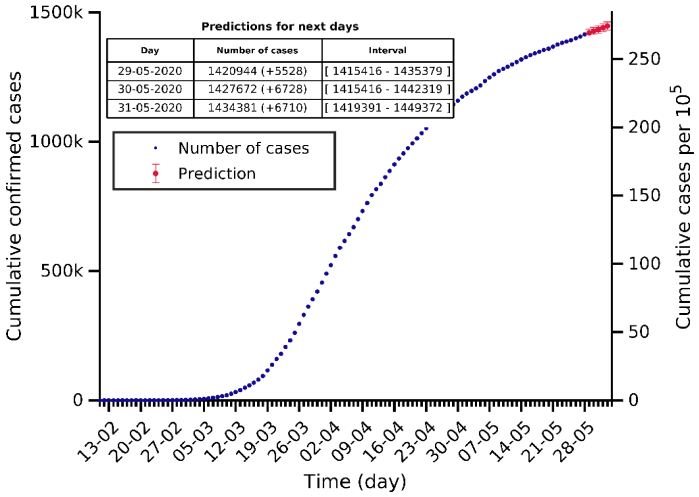


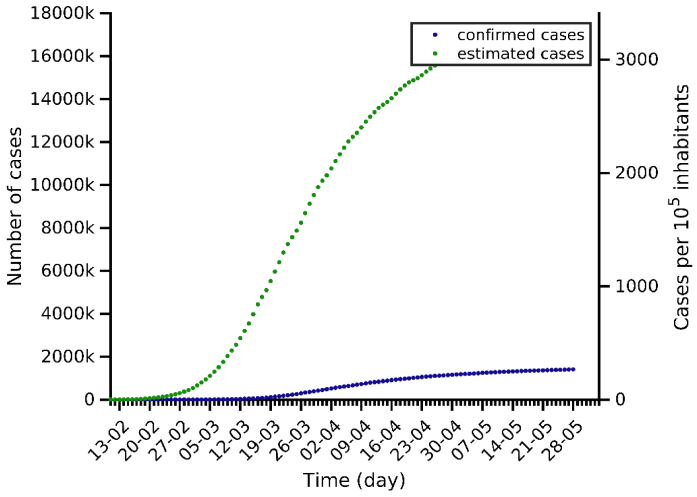
Deaths / cumulated reported cases

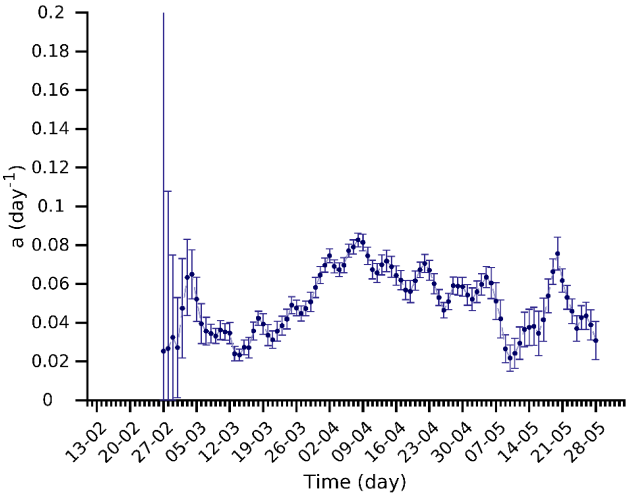
(1) Analysis and prediction of COVID-19 for EU+EFTA+UK

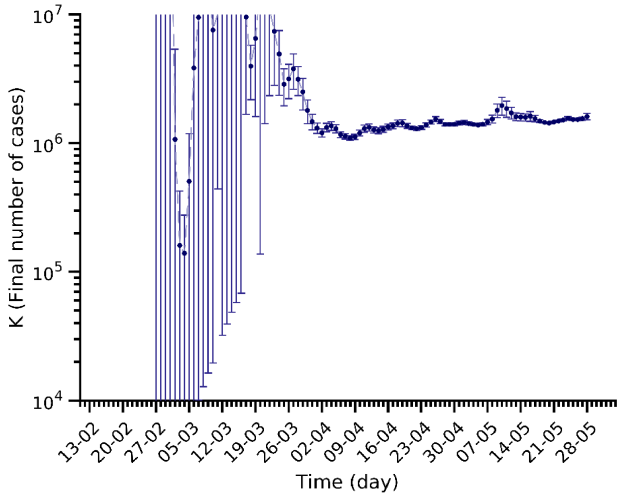
Data obtained from https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

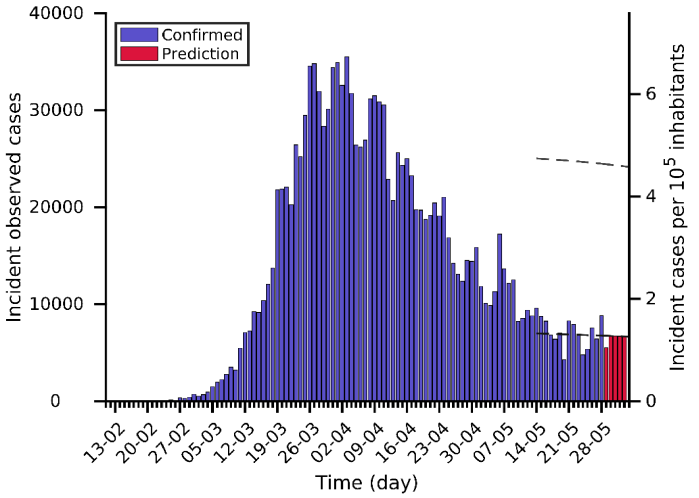
EU+EFTA+UK 28-05-2020. Population: 527.9M. Current cumulated incidence: 268/1

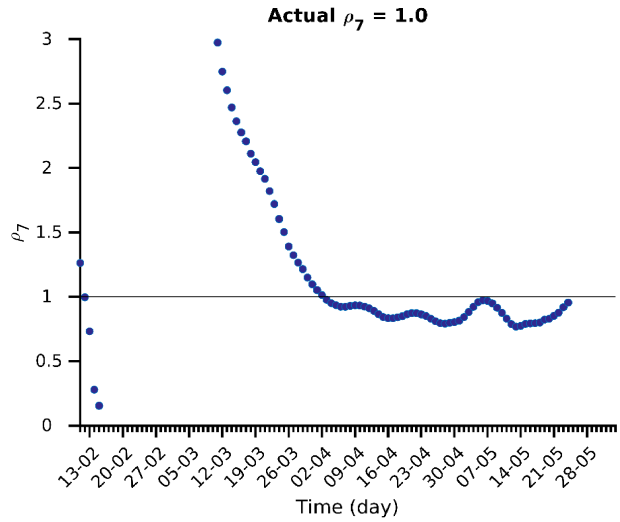


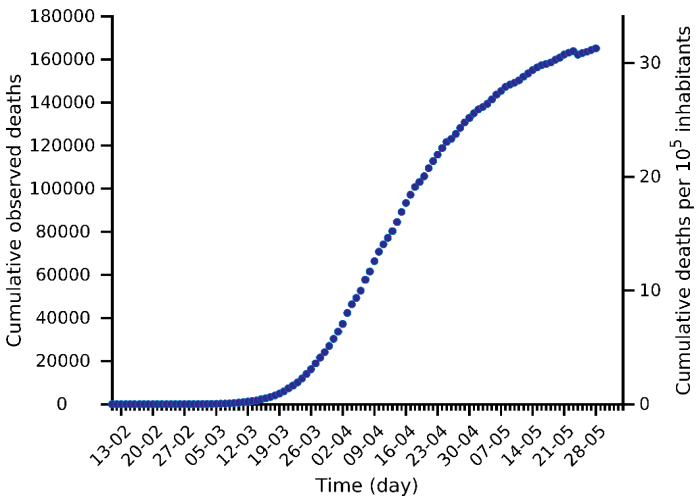


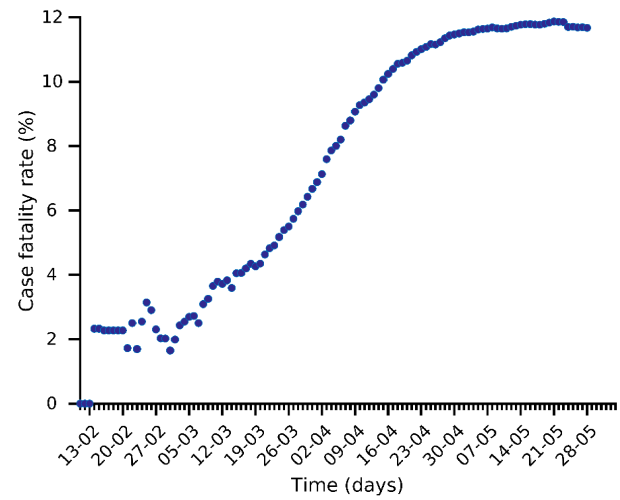




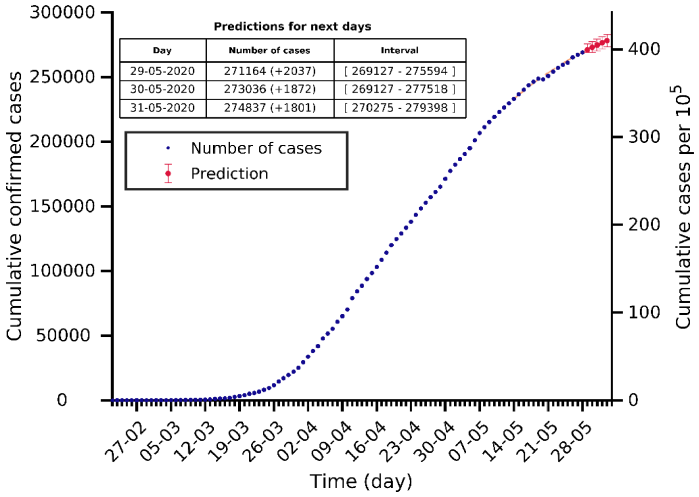


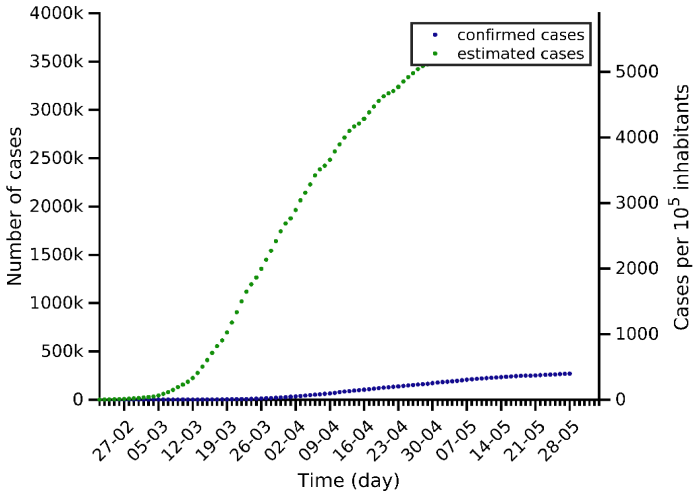


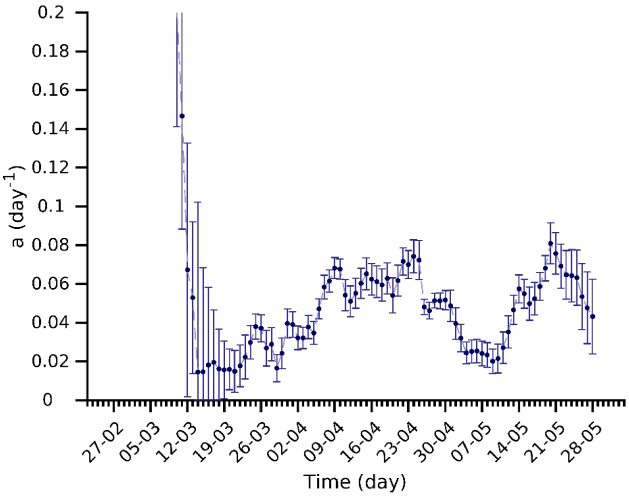


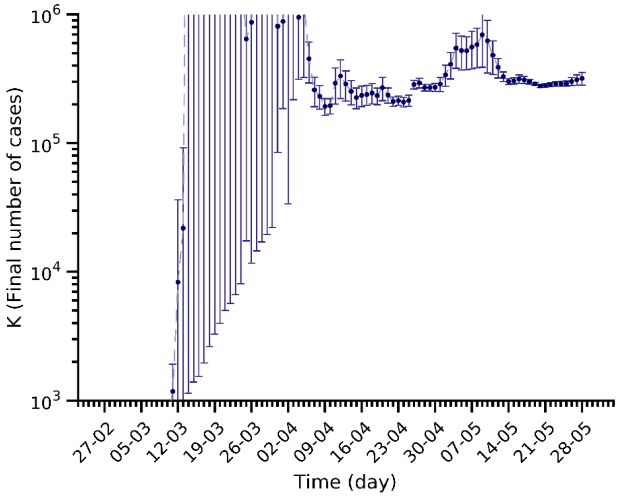


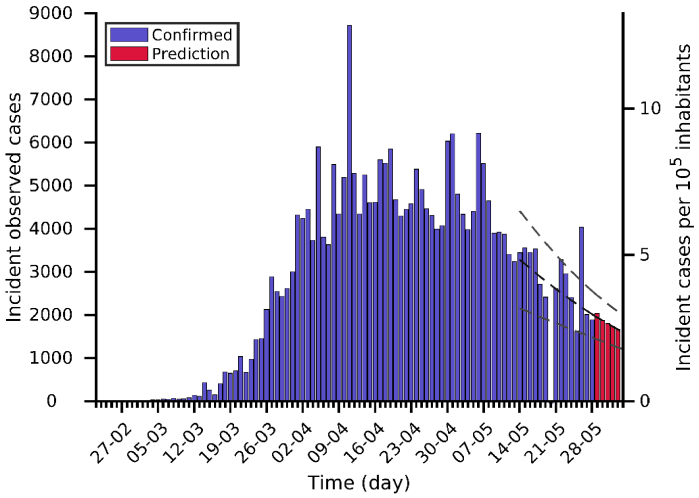
UK 28-05-2020. Population: 67.9M. Current cumulated incidence: 396/10^5

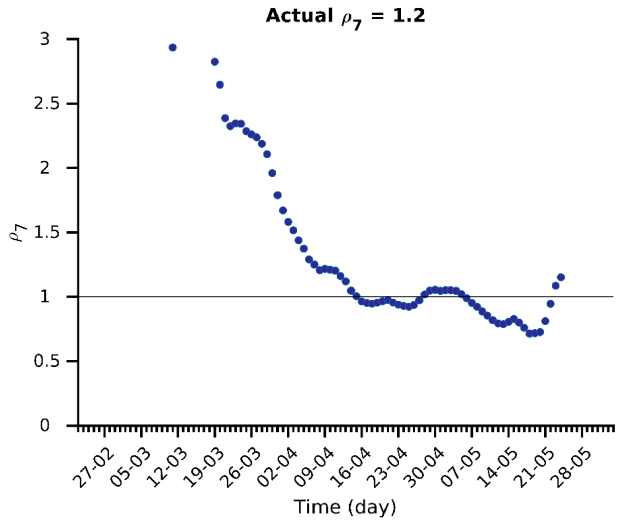


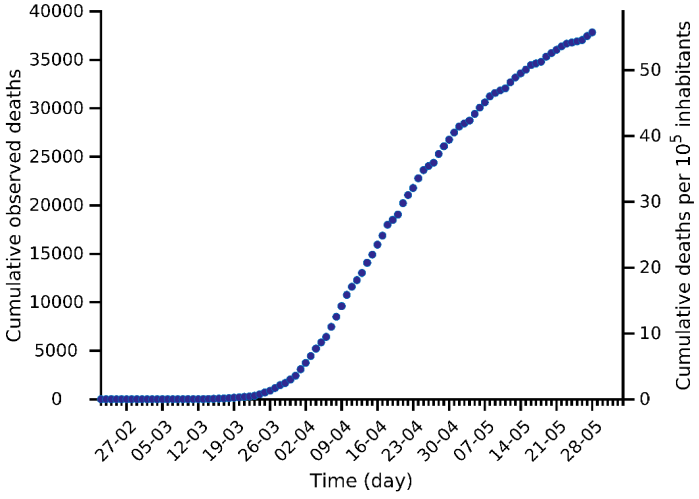


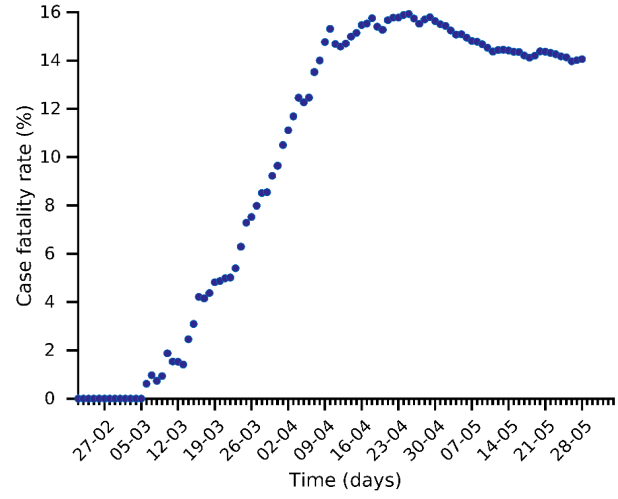




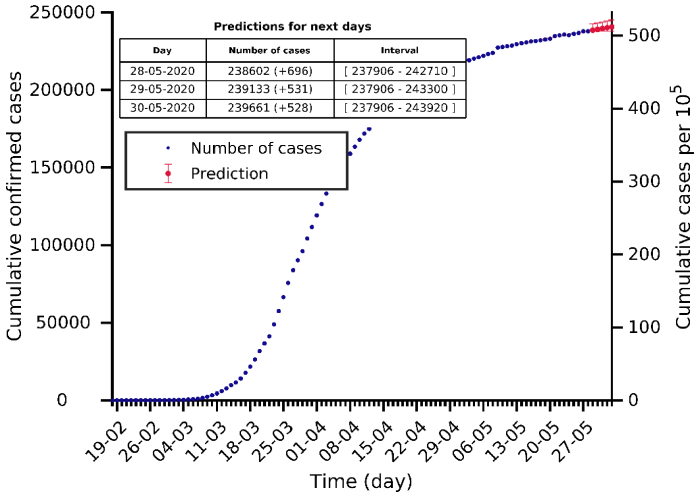


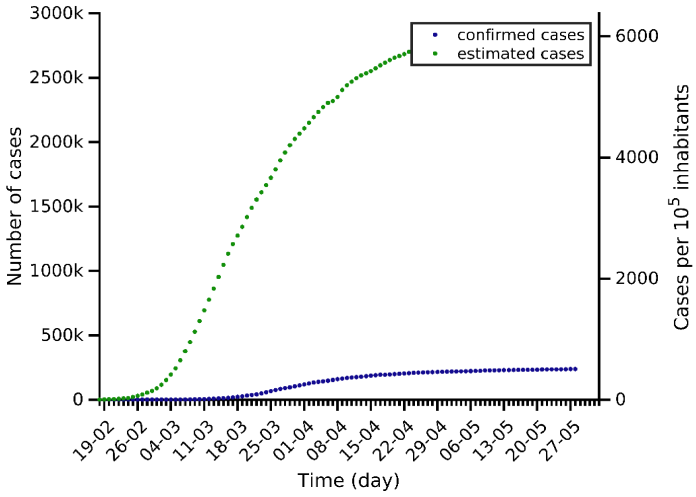


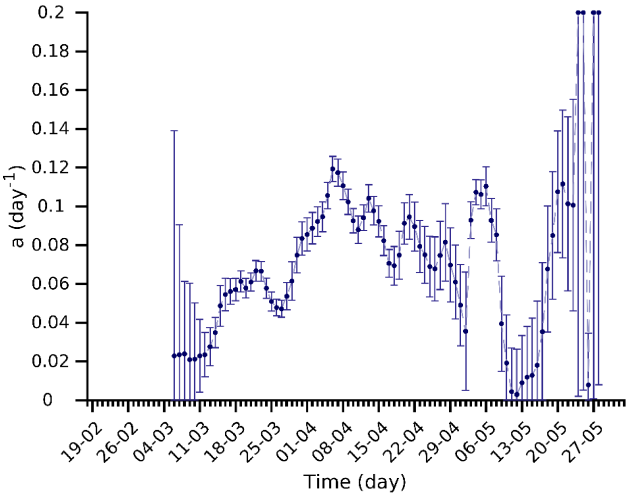


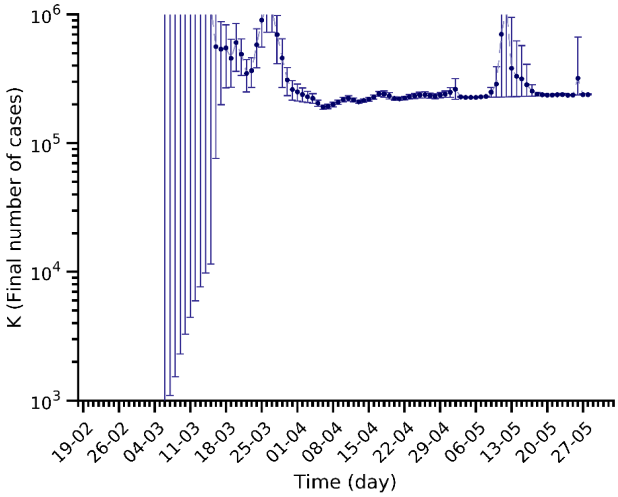


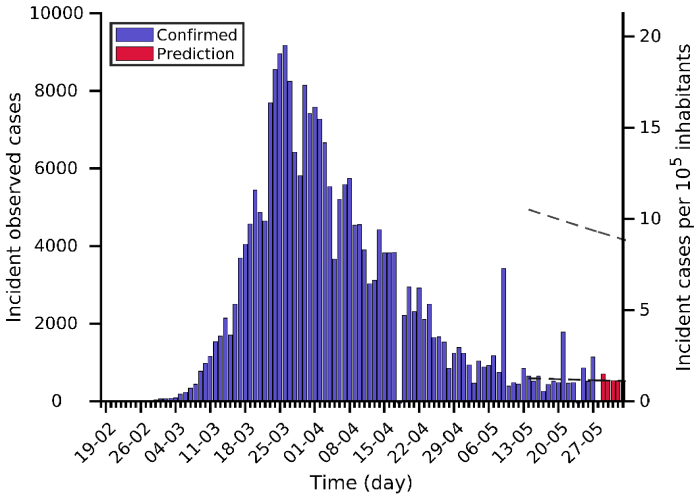
Spain 27-05-2020. Population: 47.0M. Current cumulated incidence: 506/10⁵

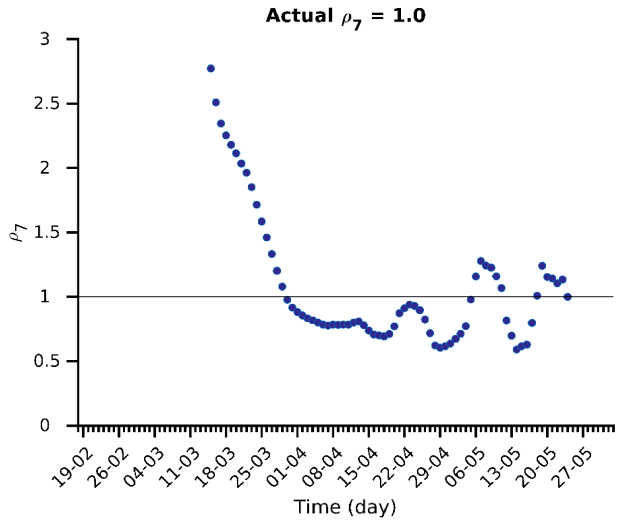


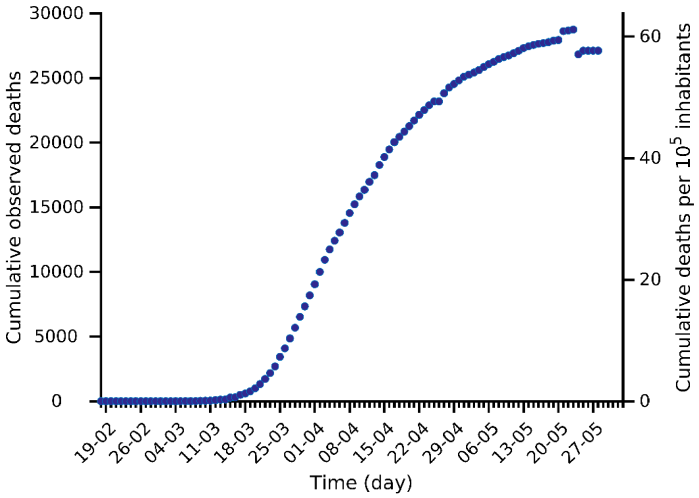


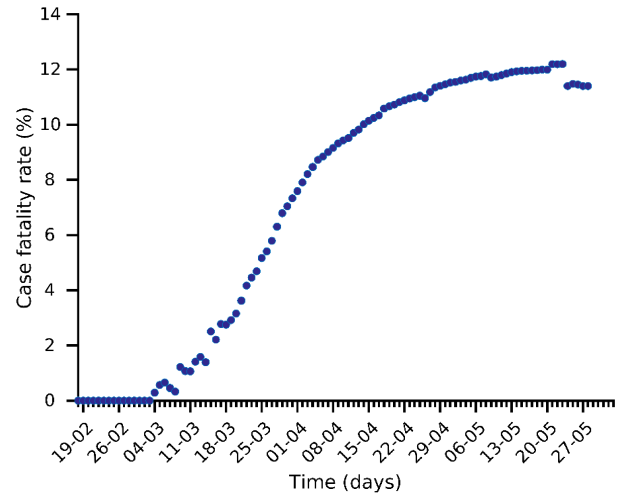




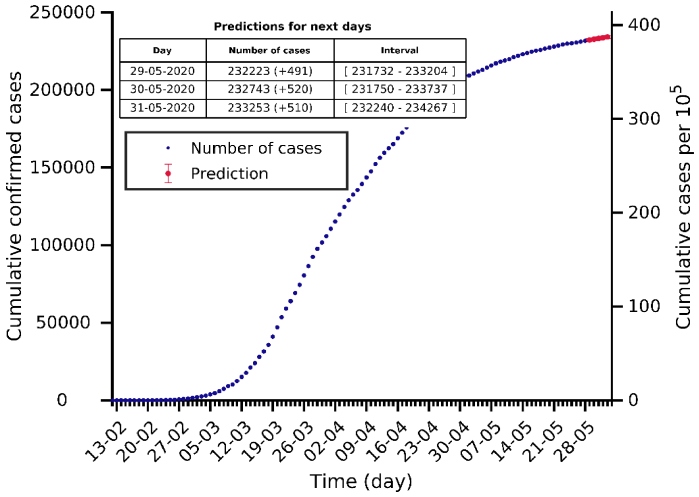


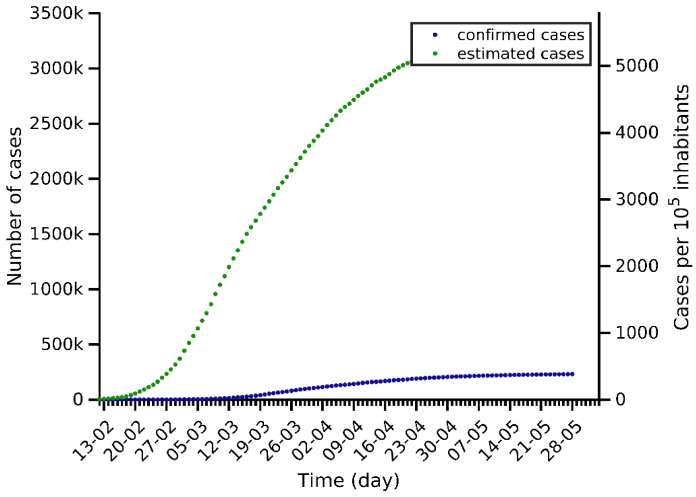


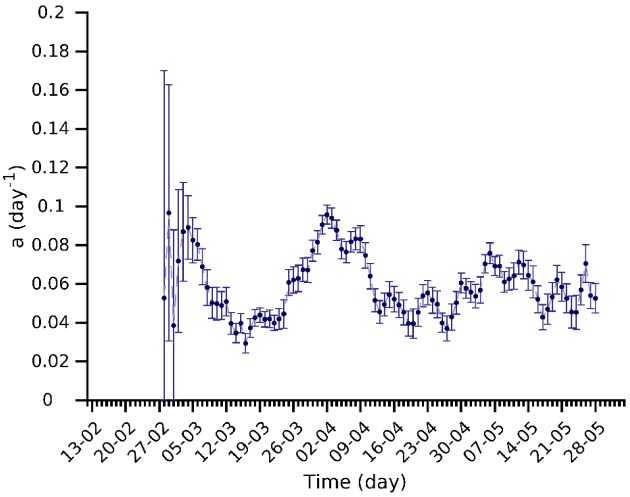


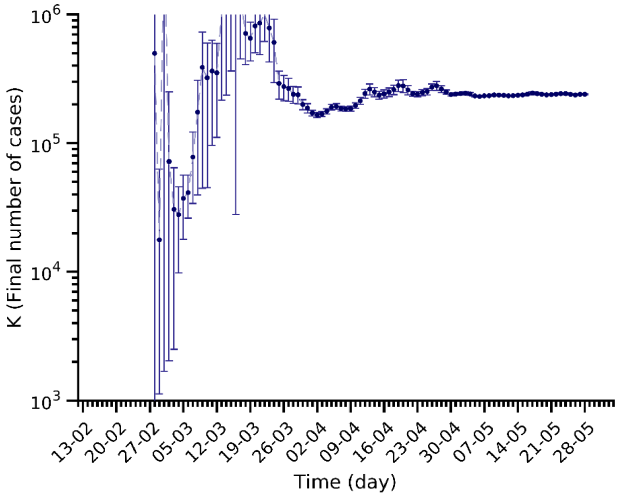


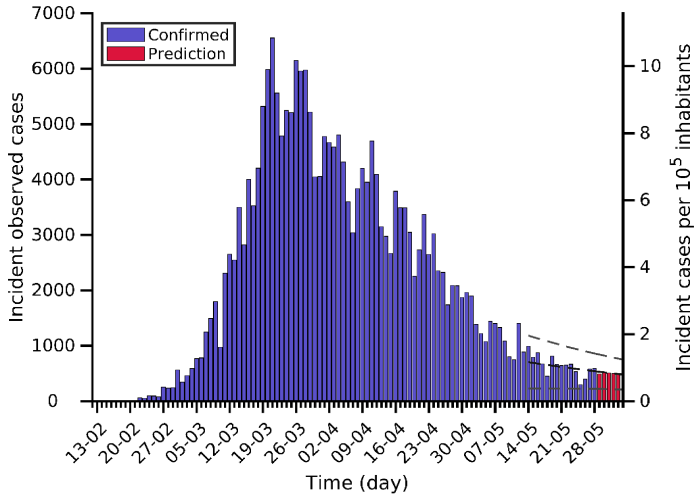
Italy 28-05-2020. Population: 60.5M. Current cumulated incidence: $383/10^5$

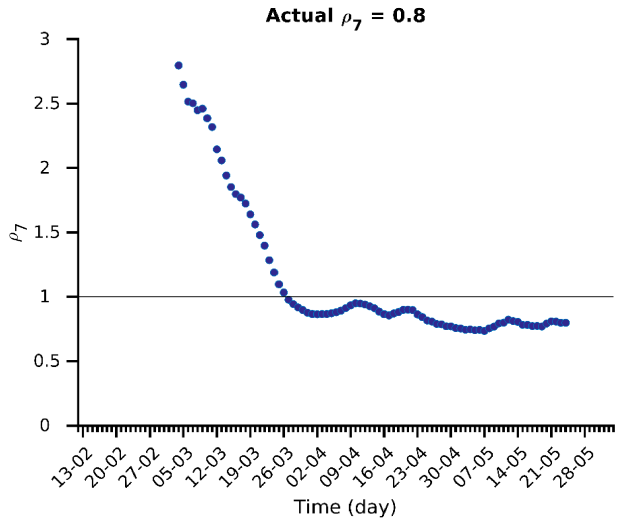


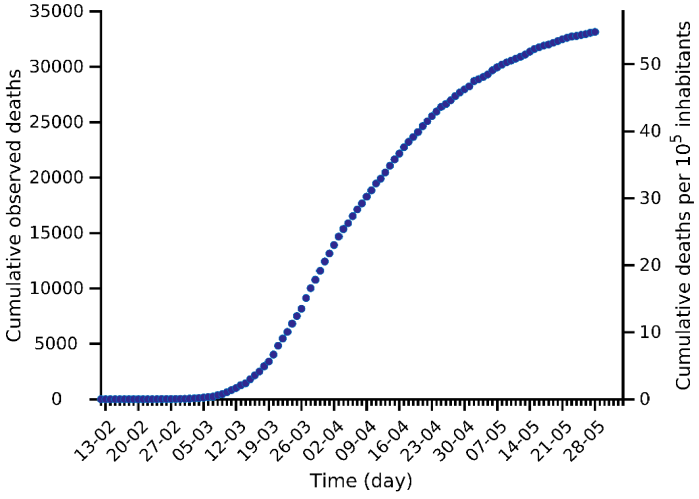


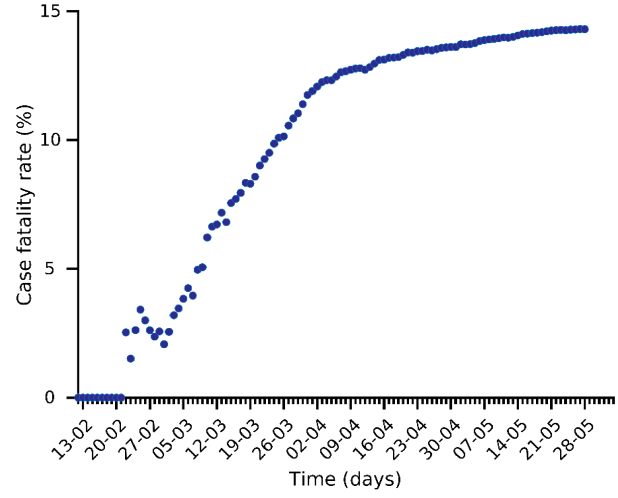




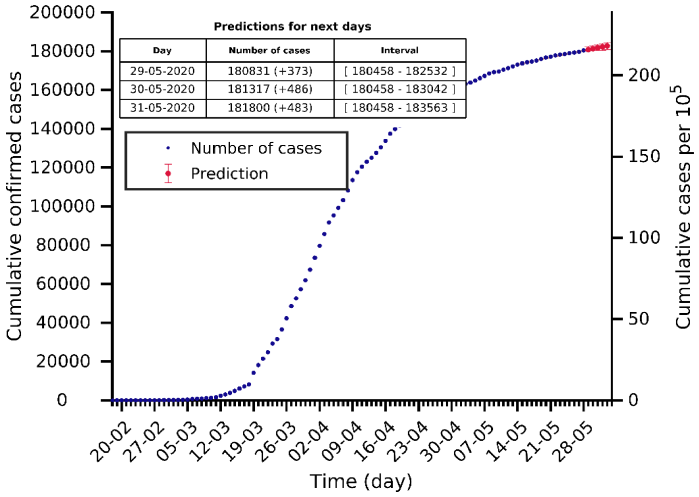


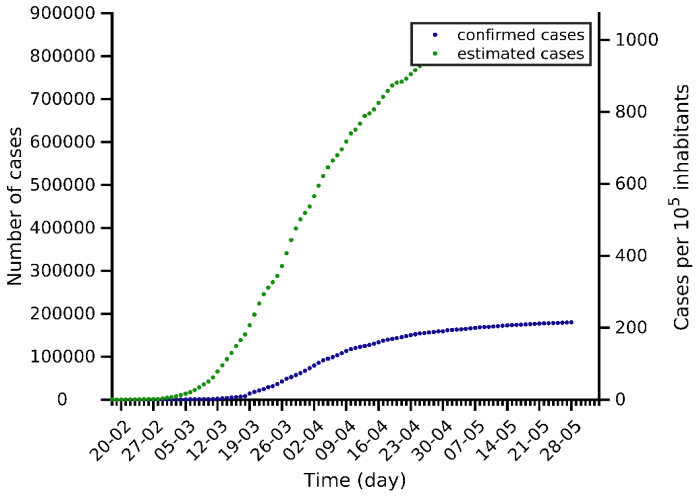


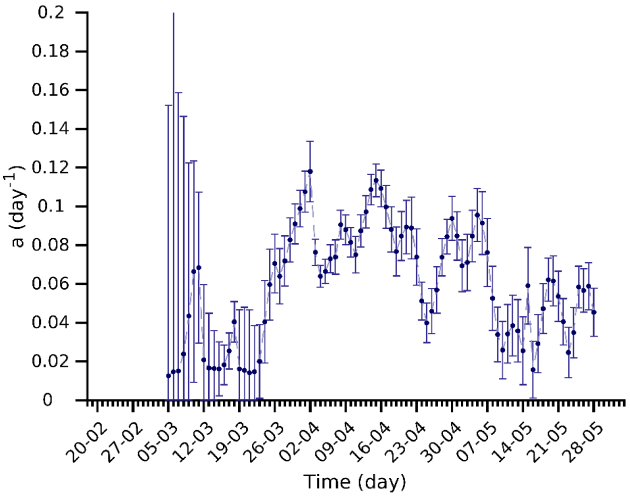


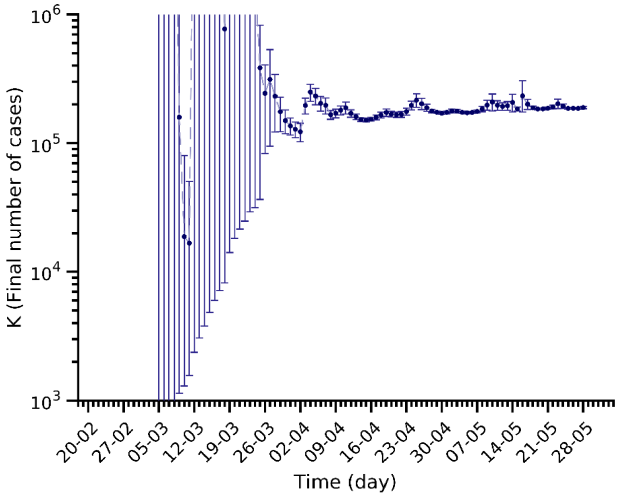


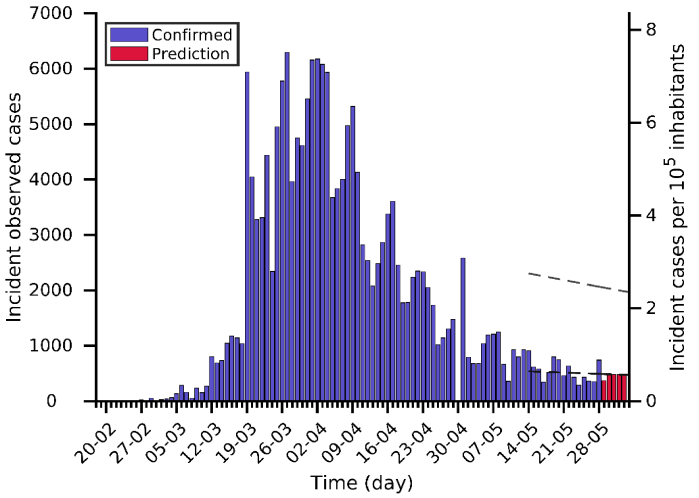
Germany 28-05-2020. Population: 83.8M. Current cumulated incidence: 215/10^5

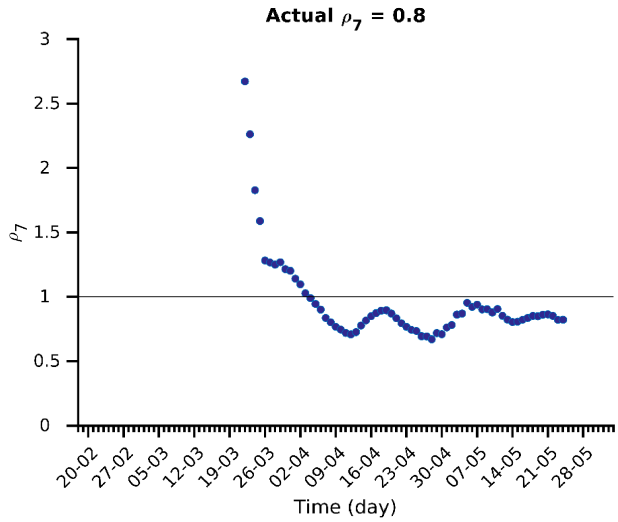


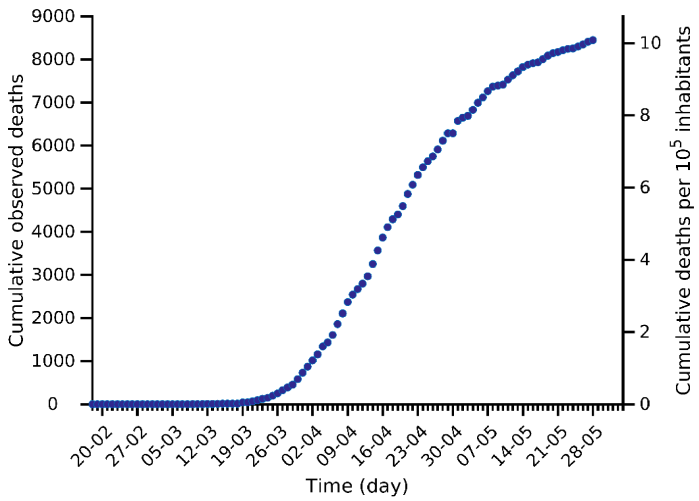


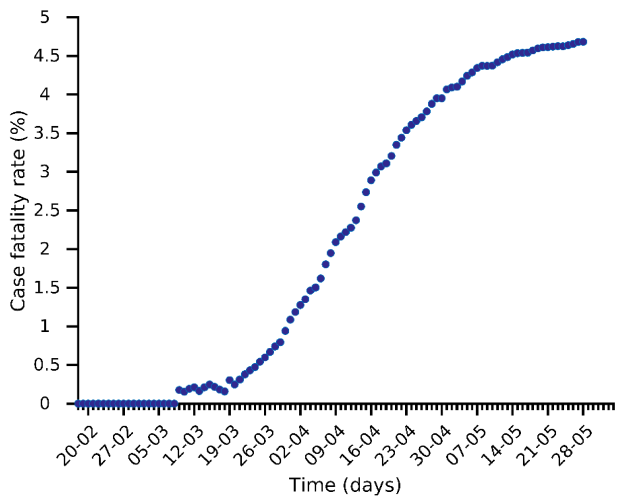




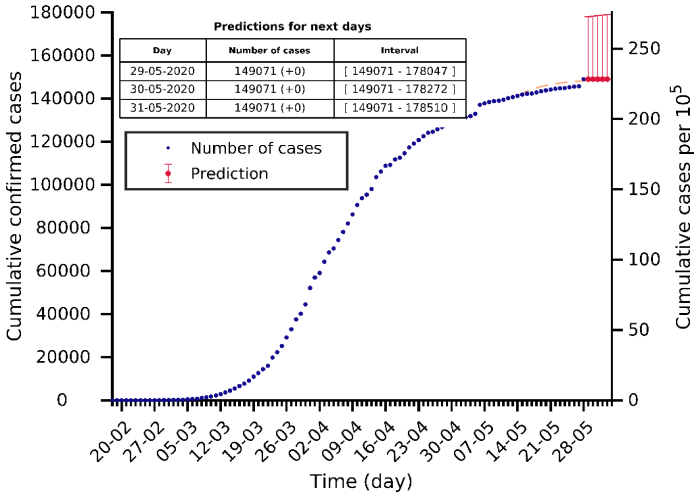


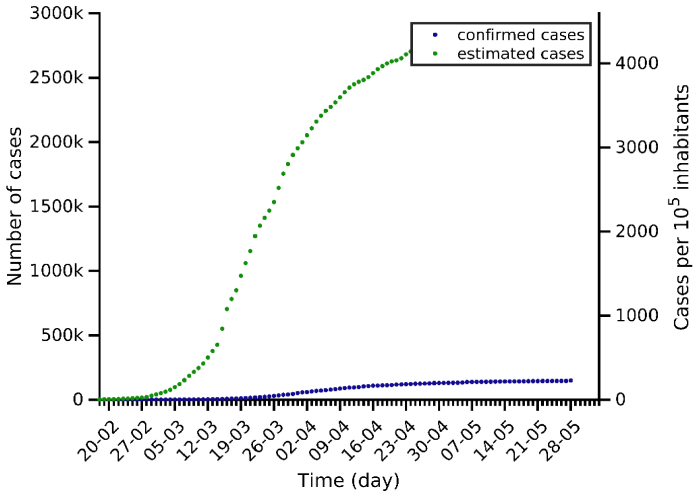


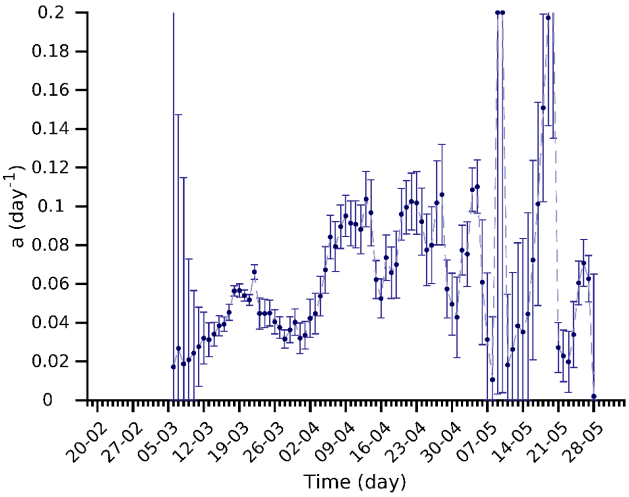


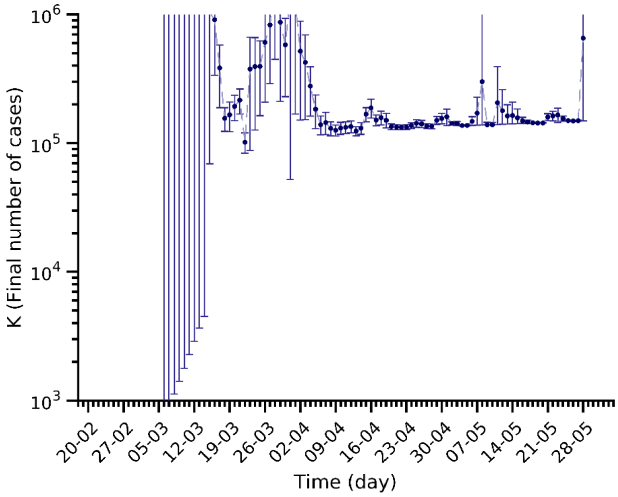


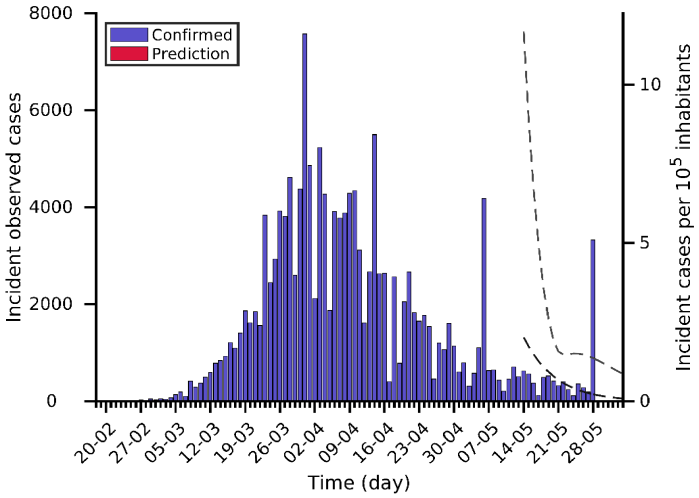
France 28-05-2020. Population: 65.3M. Current cumulated incidence: 228/10^5

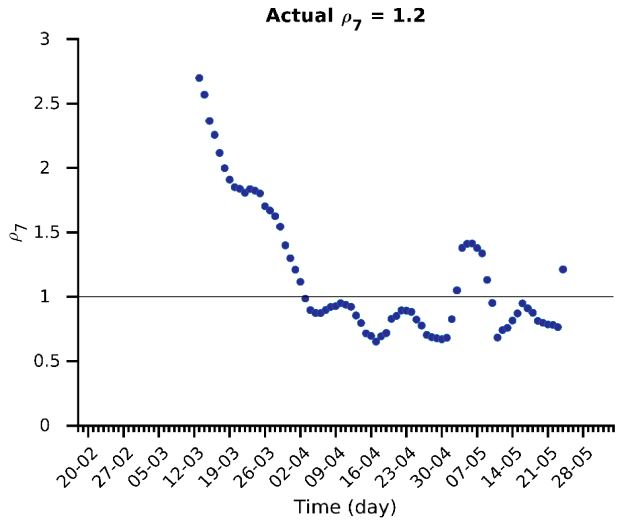


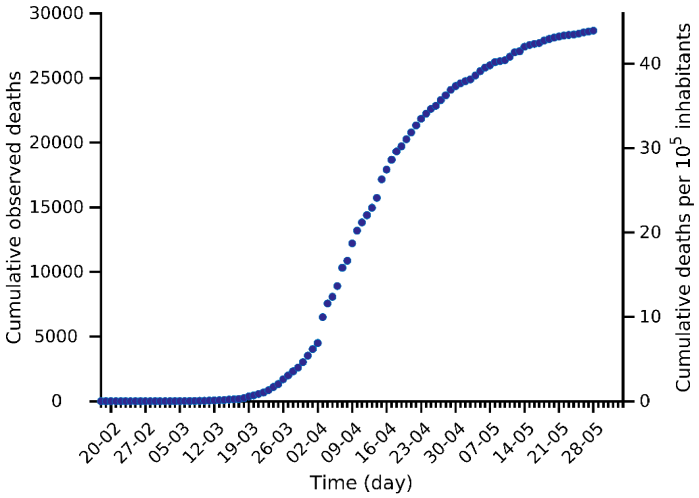


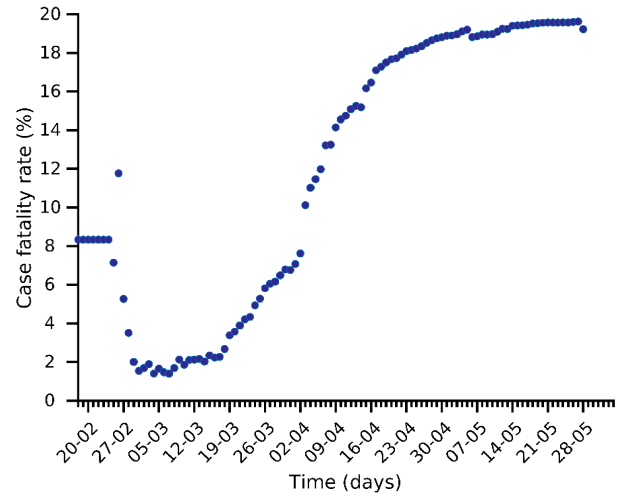




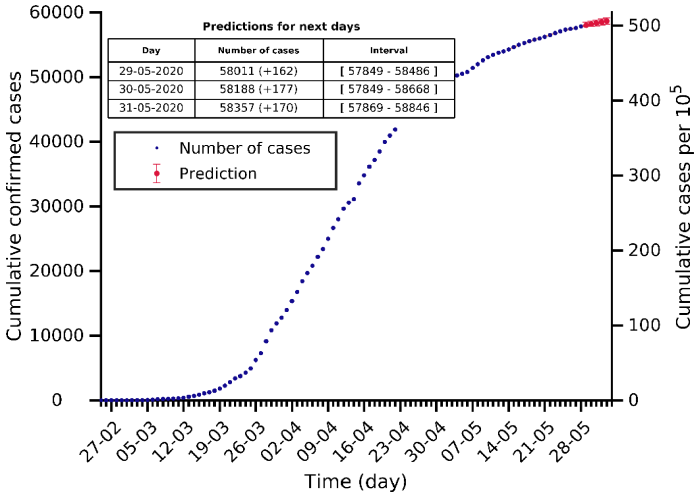


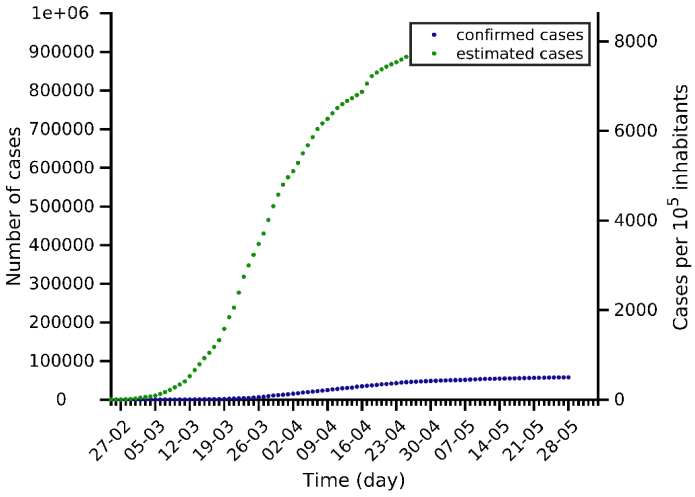


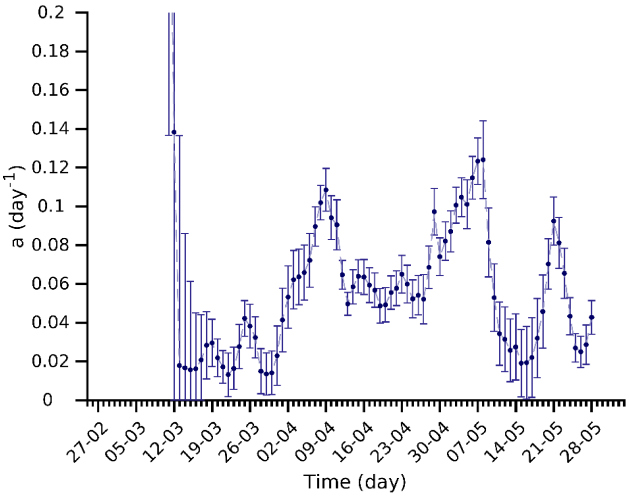


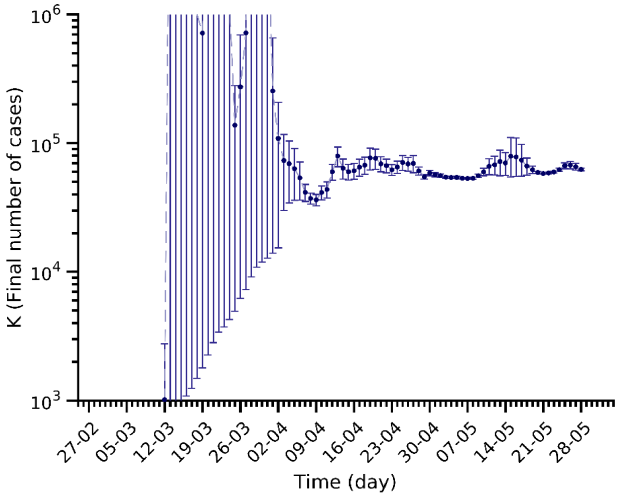


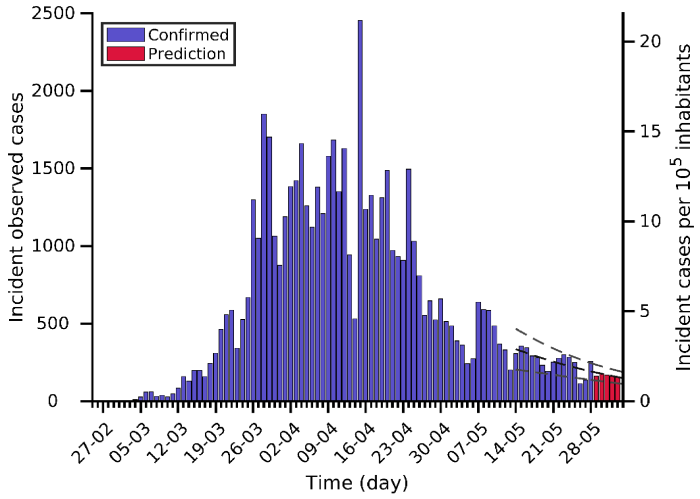
Belgium 28-05-2020. Population: 11.6M. Current cumulated incidence: 499/10^5

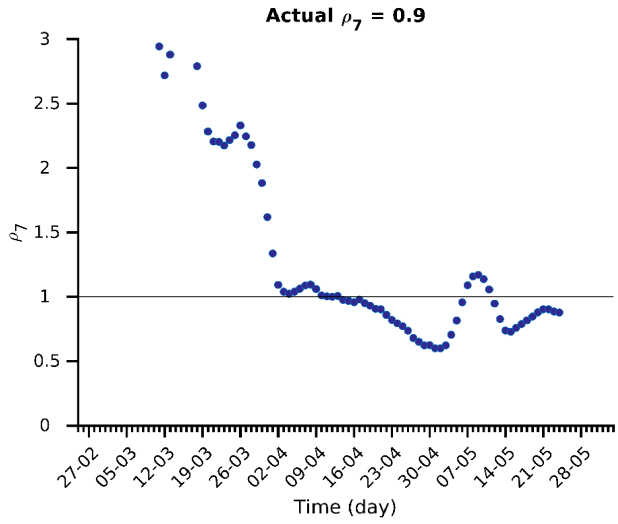


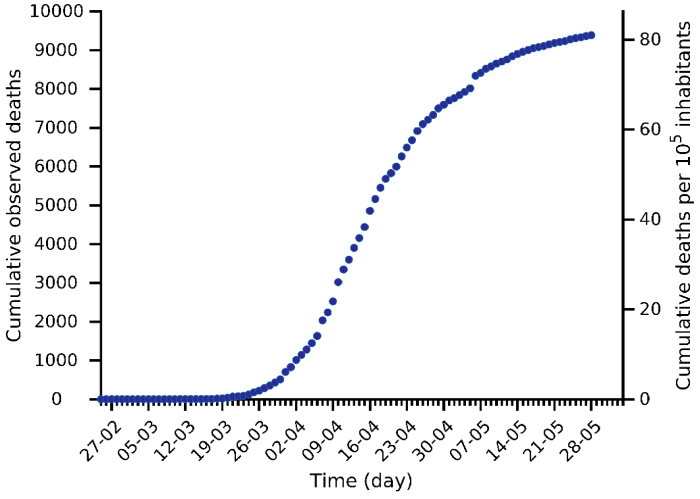


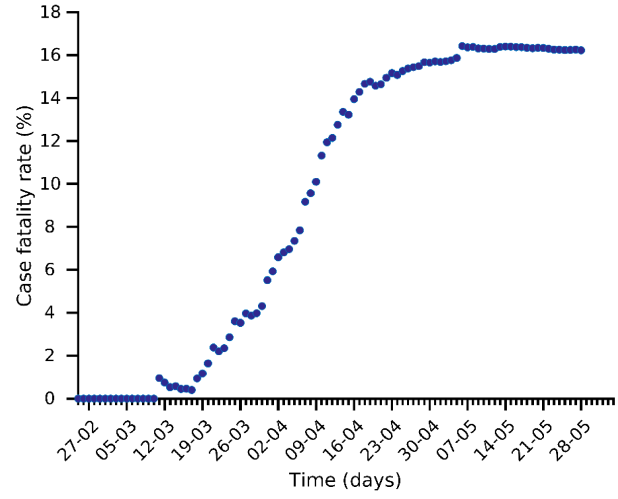




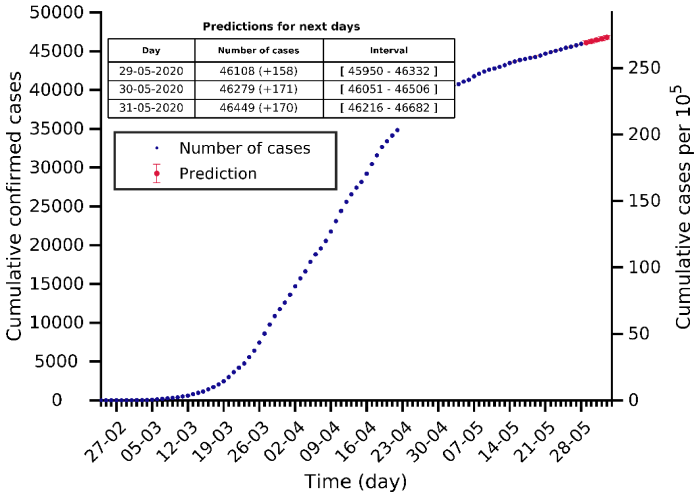


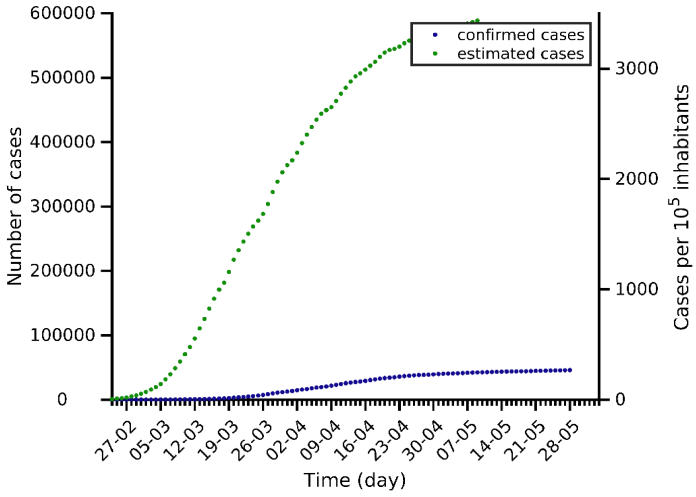


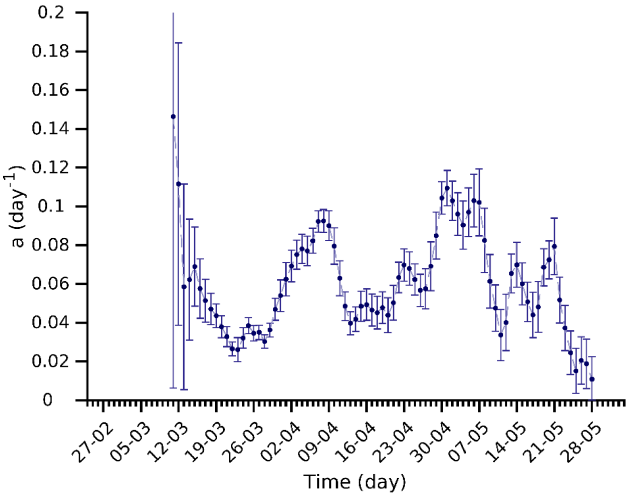


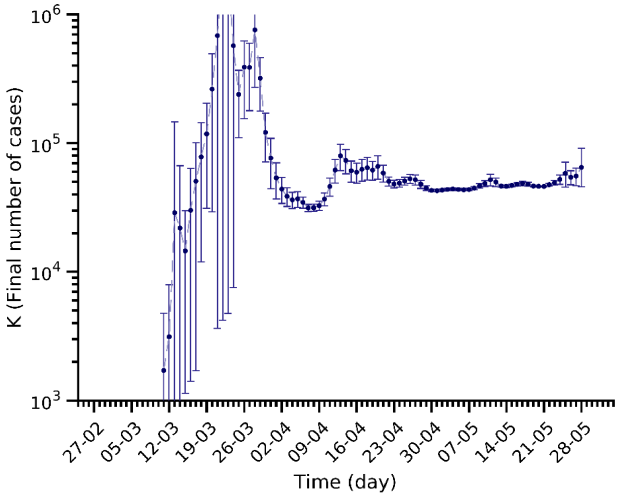


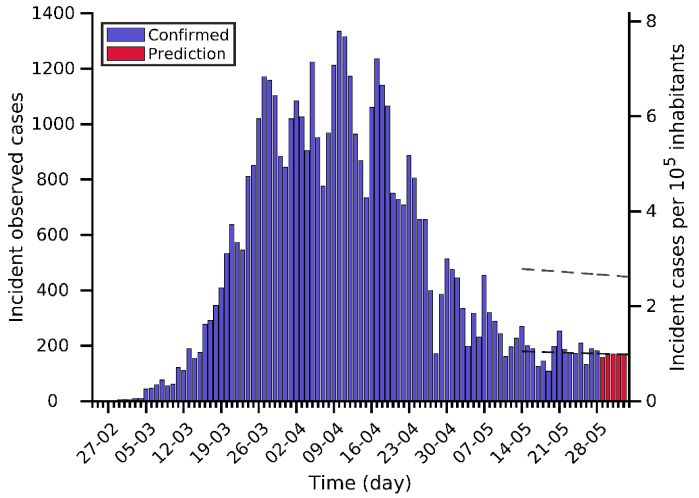
Netherlands 28-05-2020. Population: 17.1M. Current cumulated incidence: 268/10^5

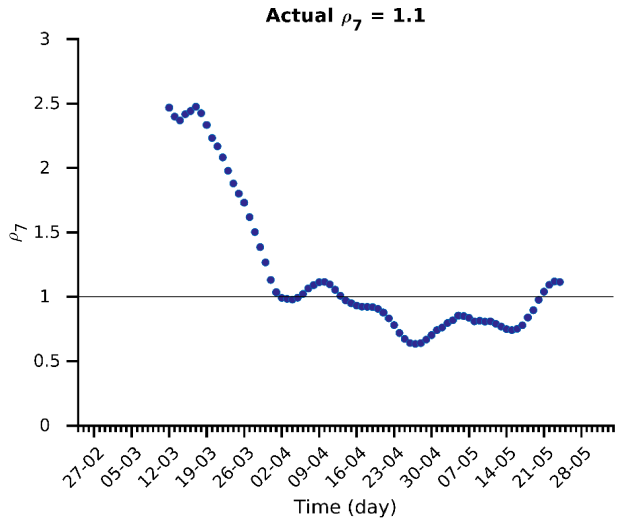


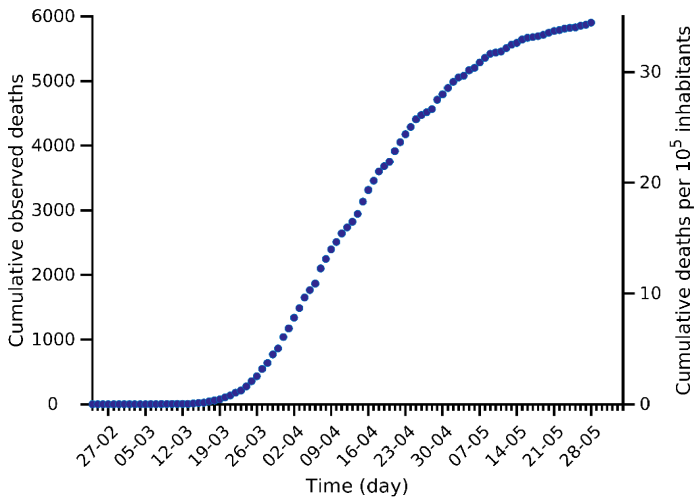


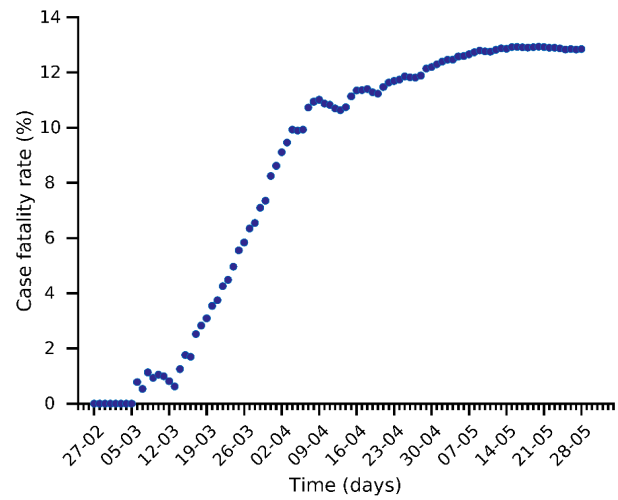




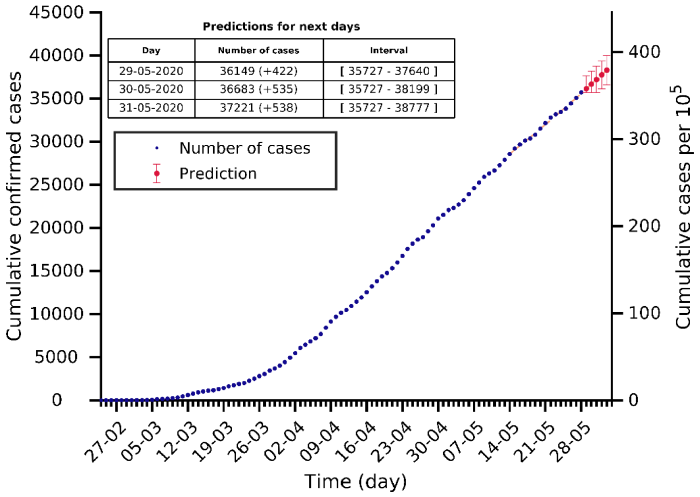


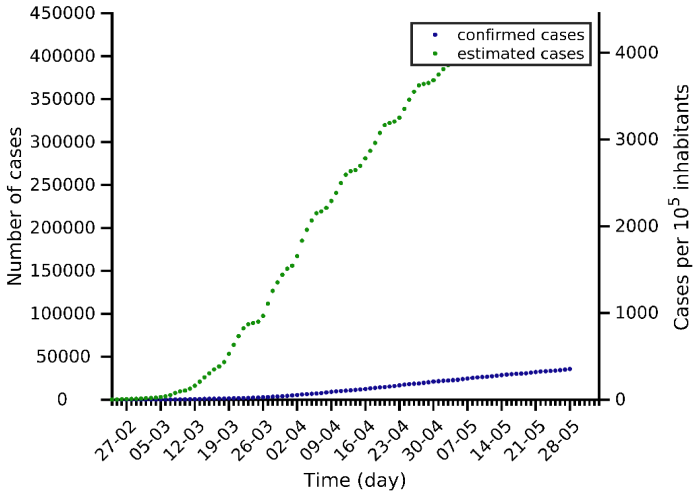


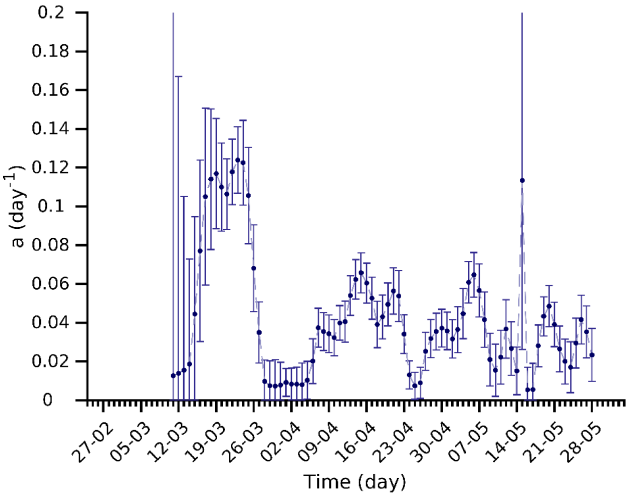


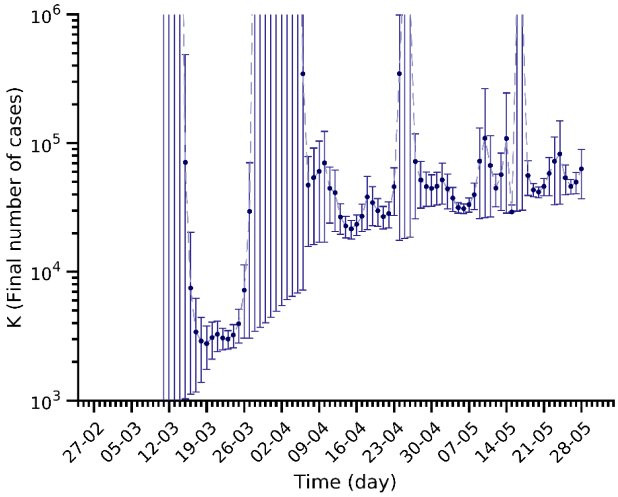


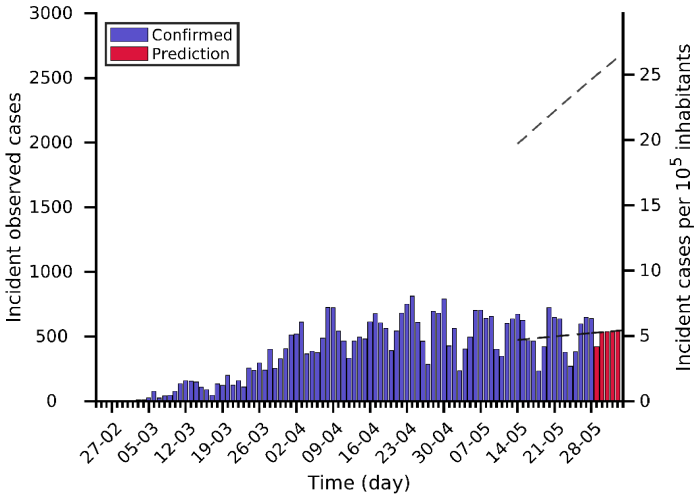
Sweden 28-05-2020. Population: 10.1M. Current cumulated incidence: 354/10^5

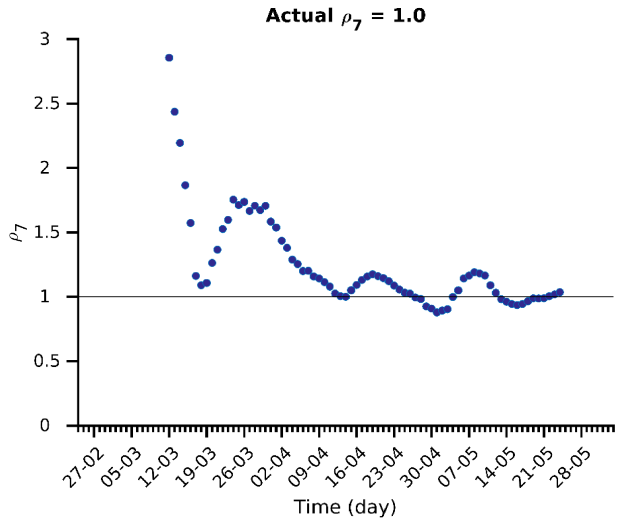


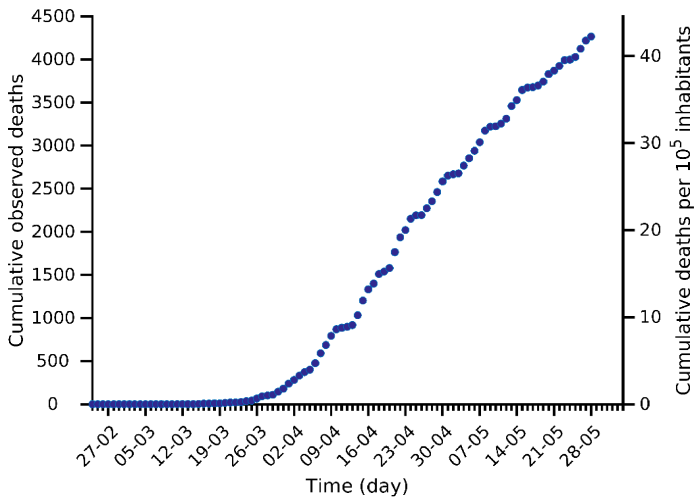


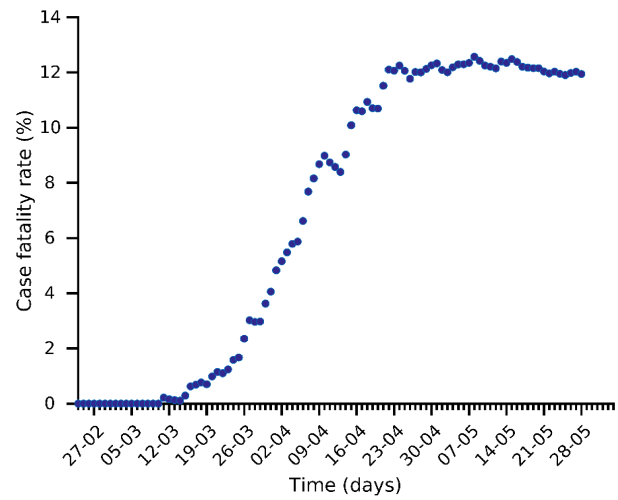




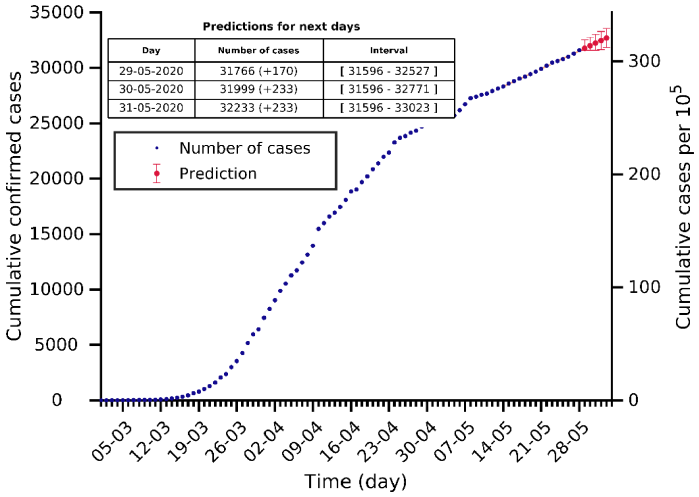


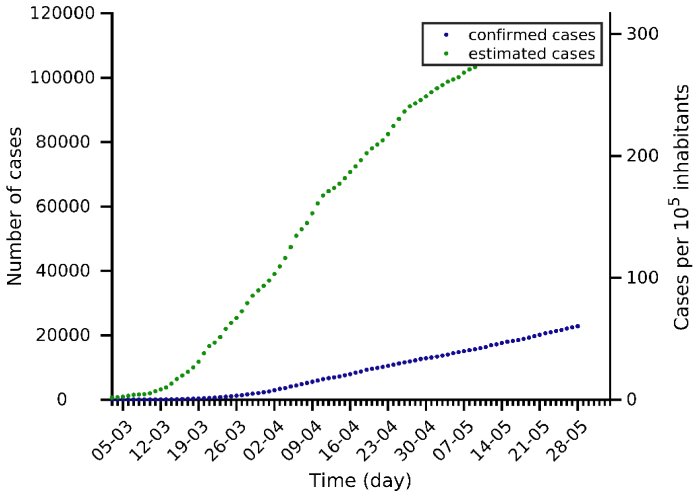


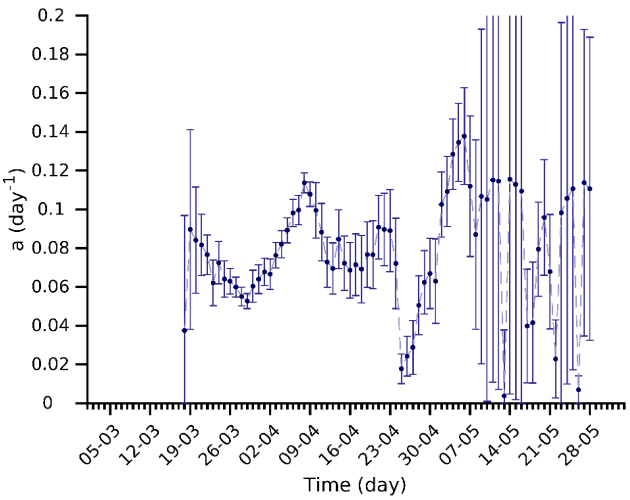


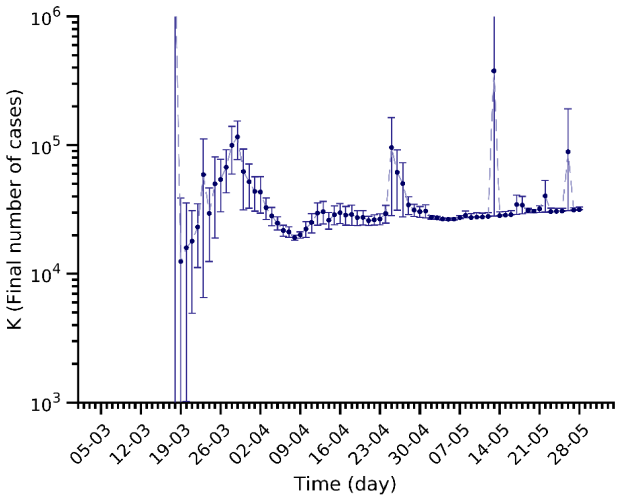


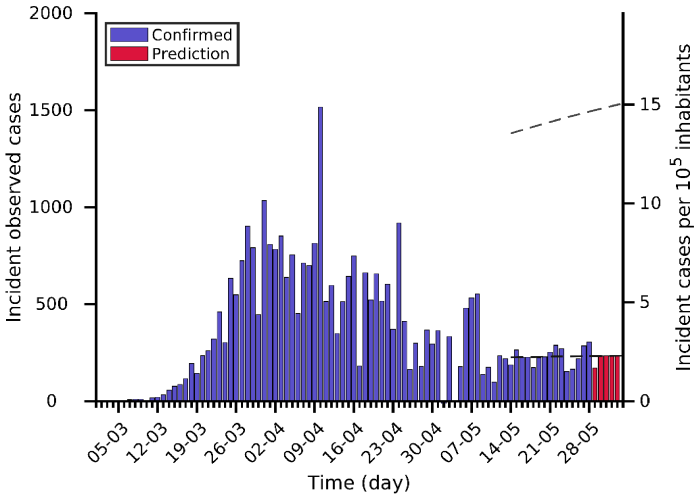
Portugal 28-05-2020. Population: 10.2M. Current cumulated incidence: 310/10⁵

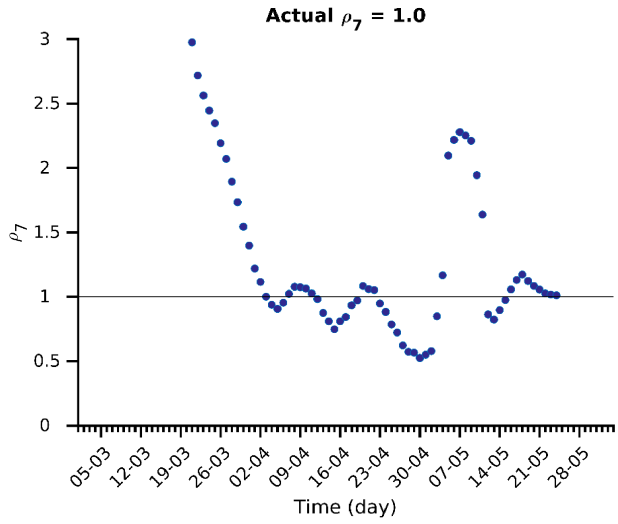


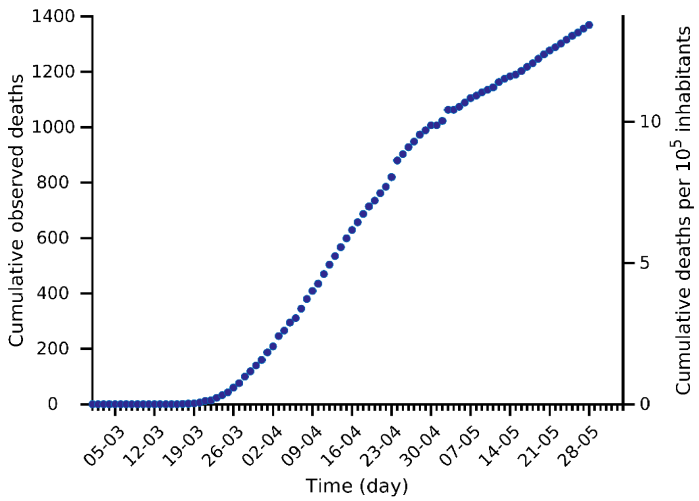


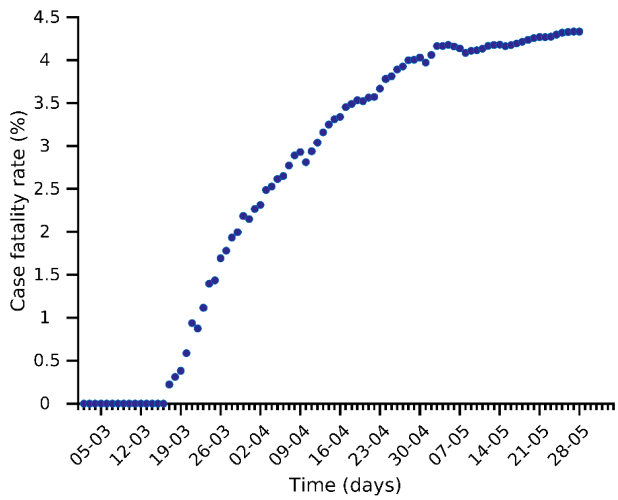




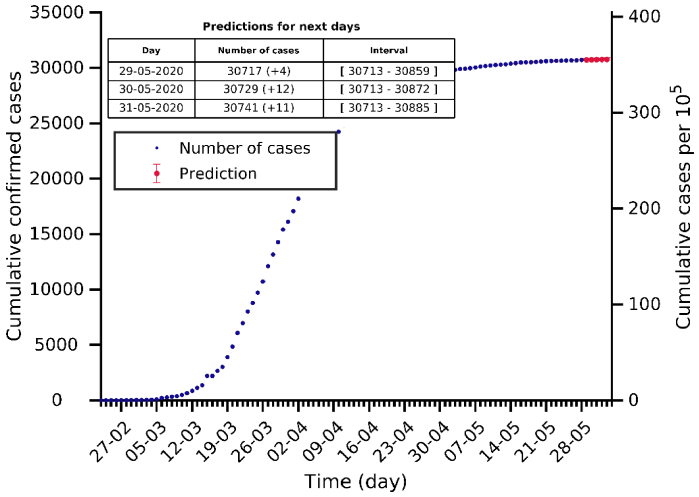


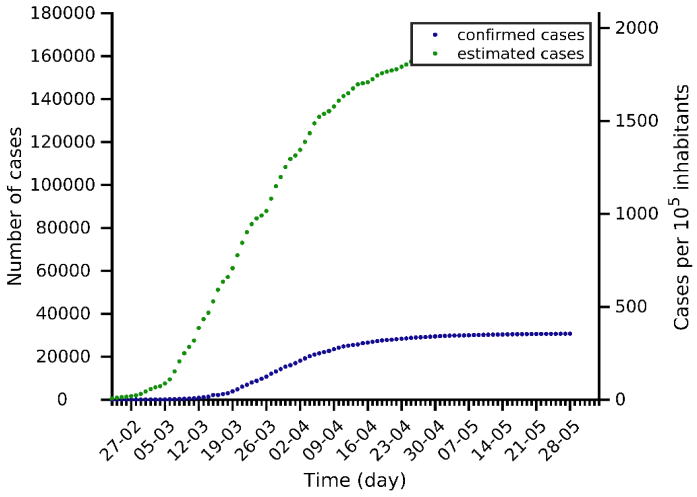


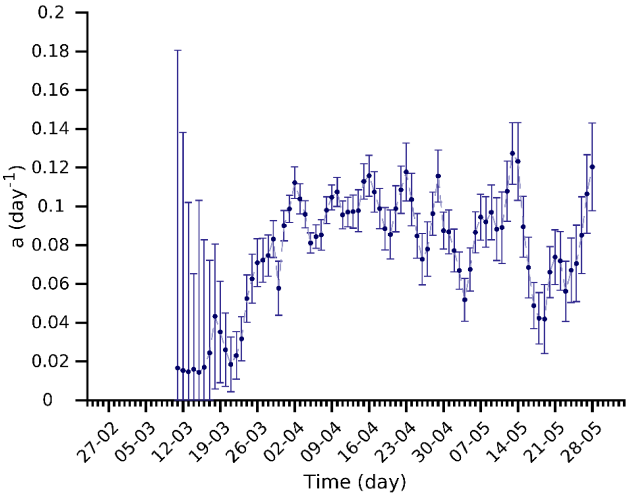


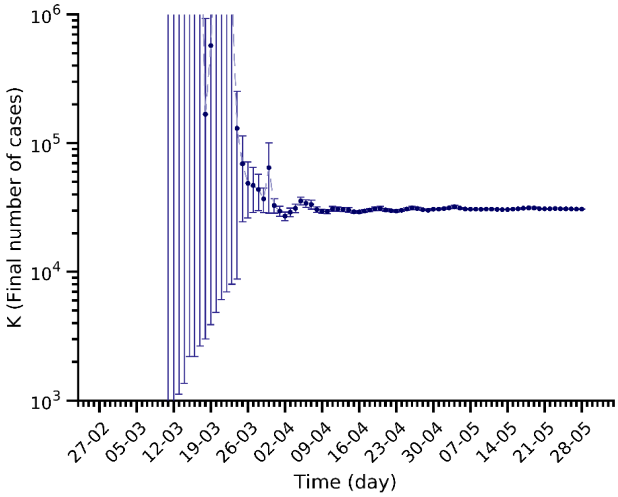


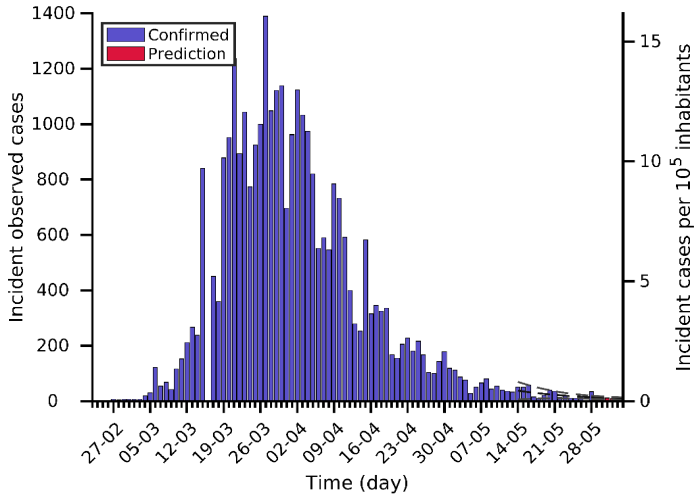
Switzerland 28-05-2020. Population: 8.7M. Current cumulated incidence: $355/10^5$

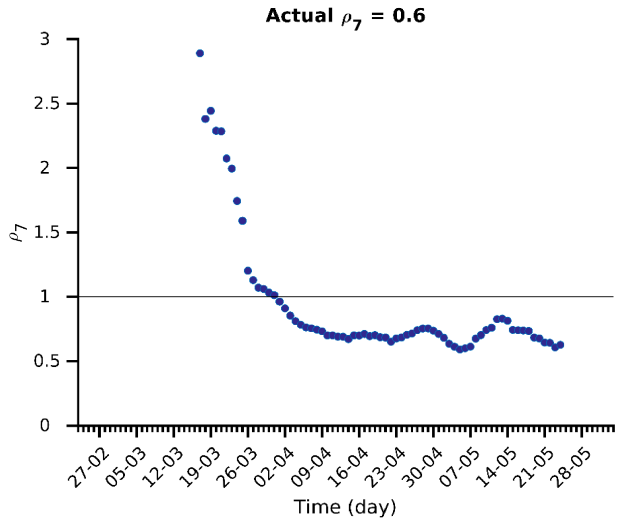


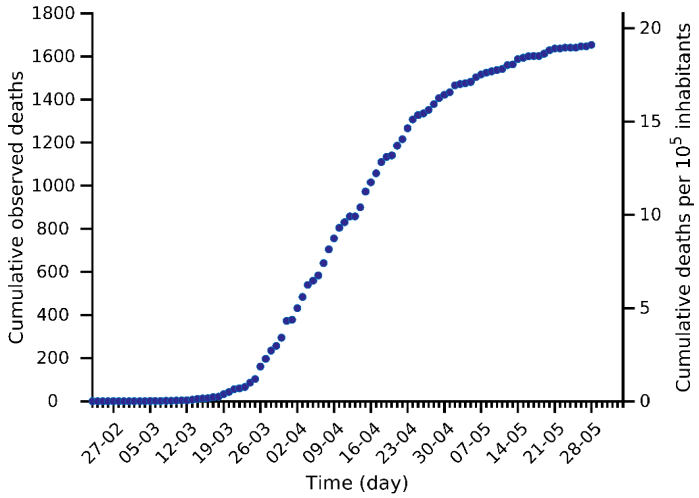


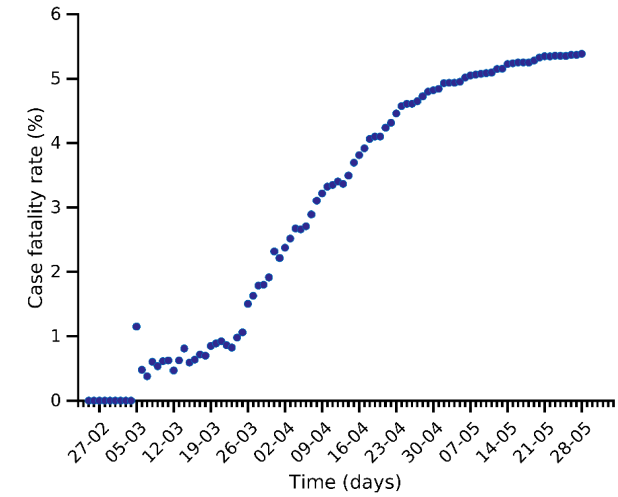




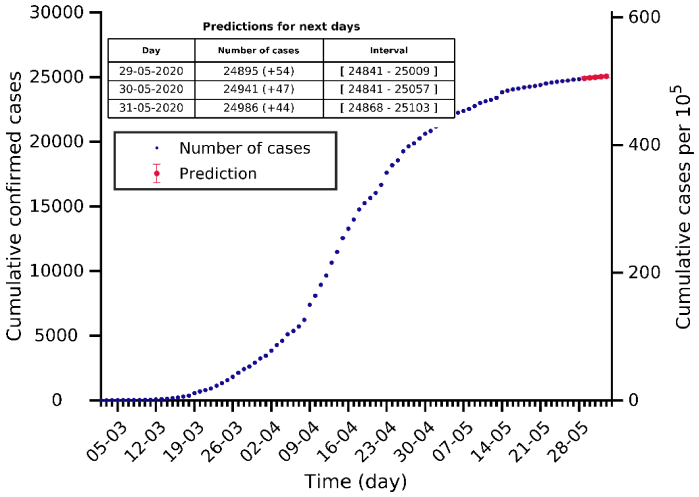


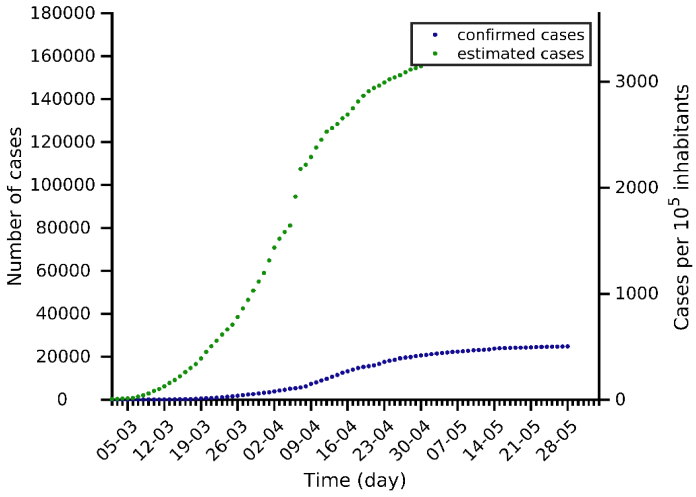


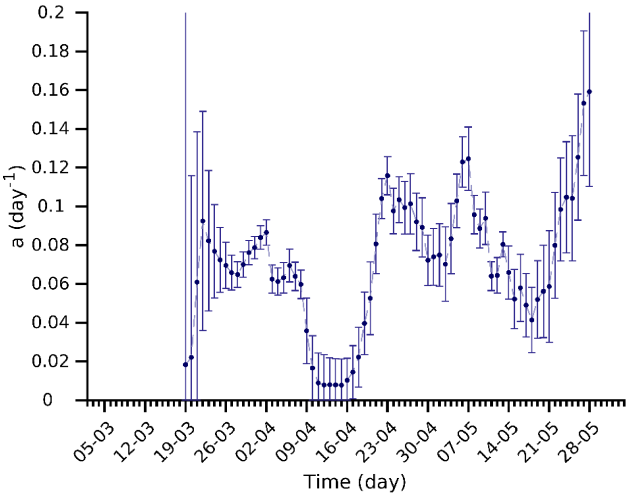


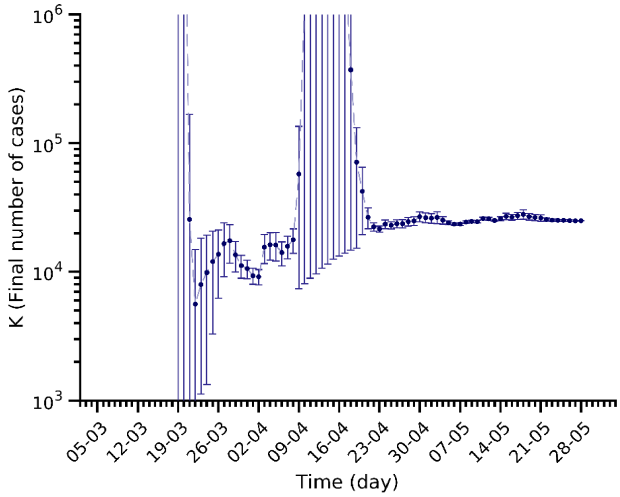


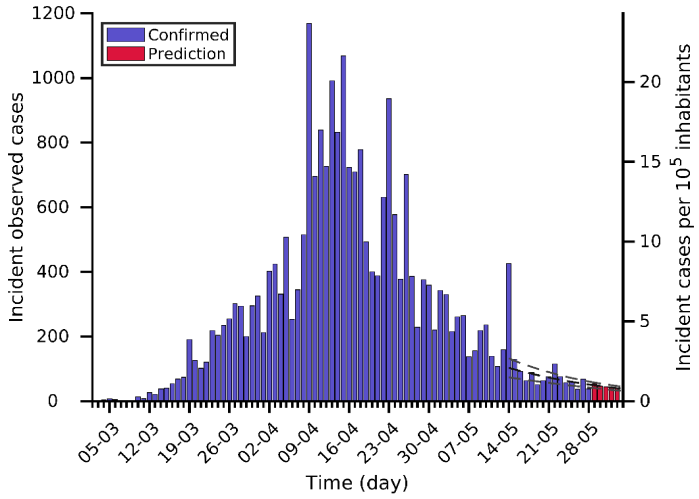
Ireland 28-05-2020. Population: 4.9M. Current cumulated incidence: 503/10^5

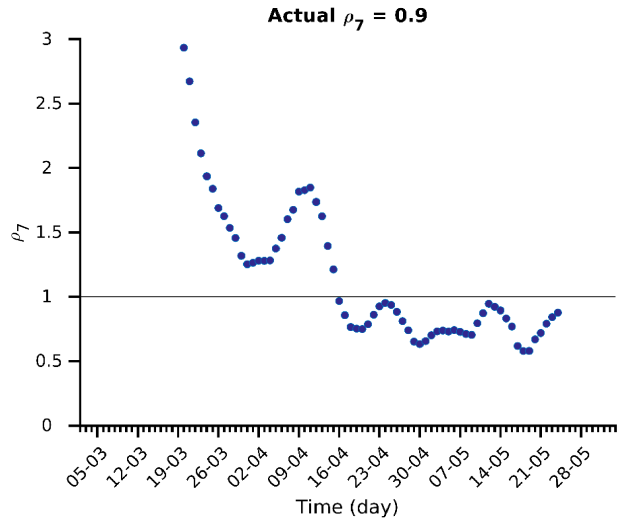


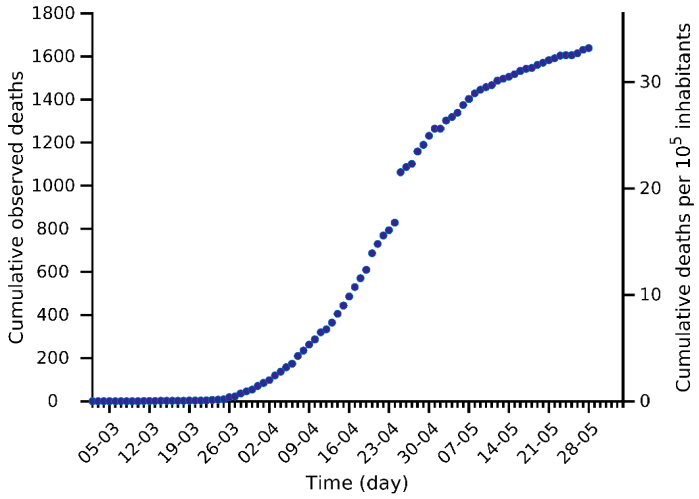


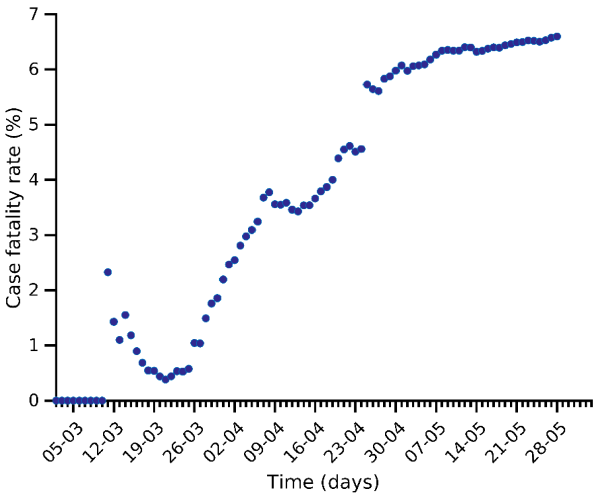




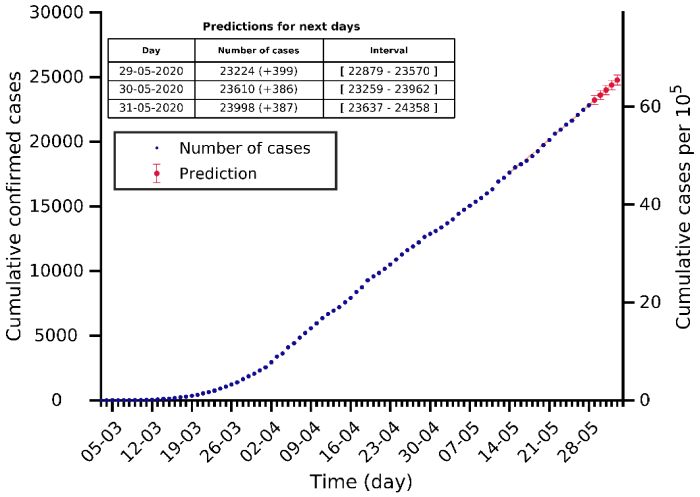


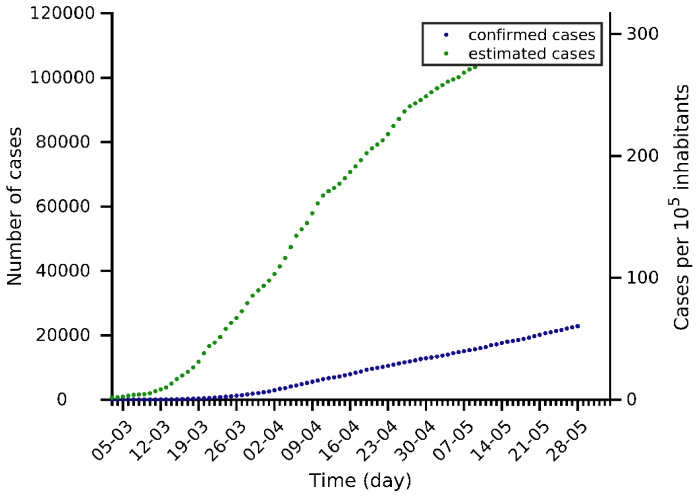


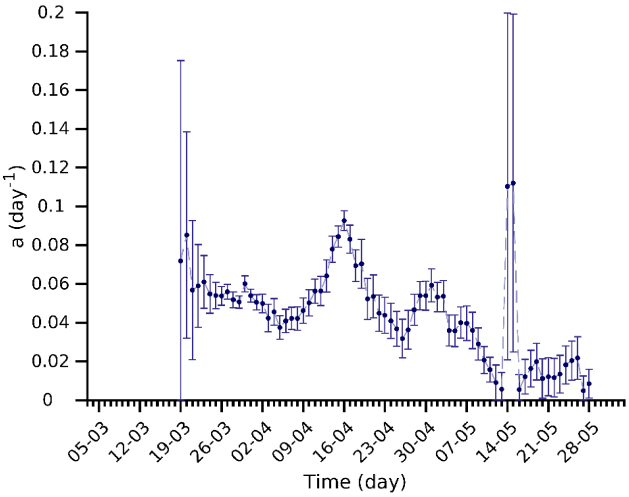


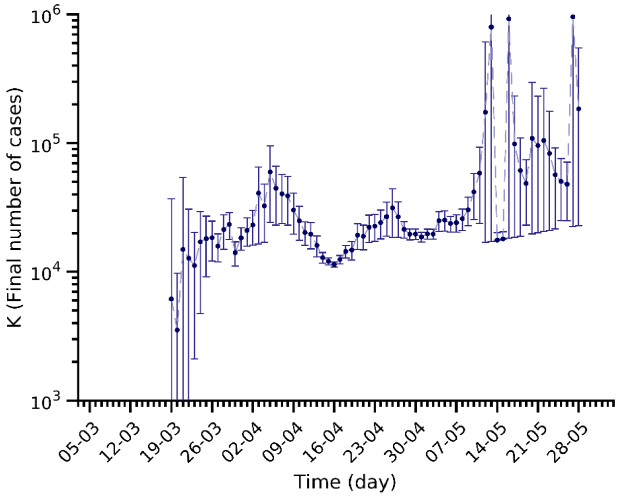


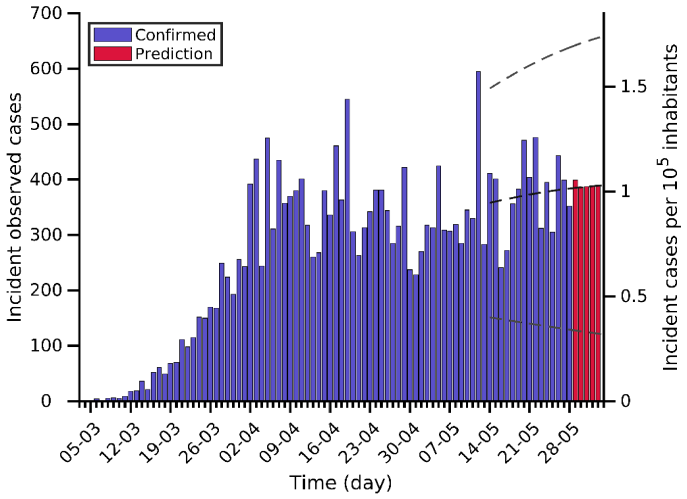
Poland 28-05-2020. Population: 37.8M. Current cumulated incidence: $60/10^5$

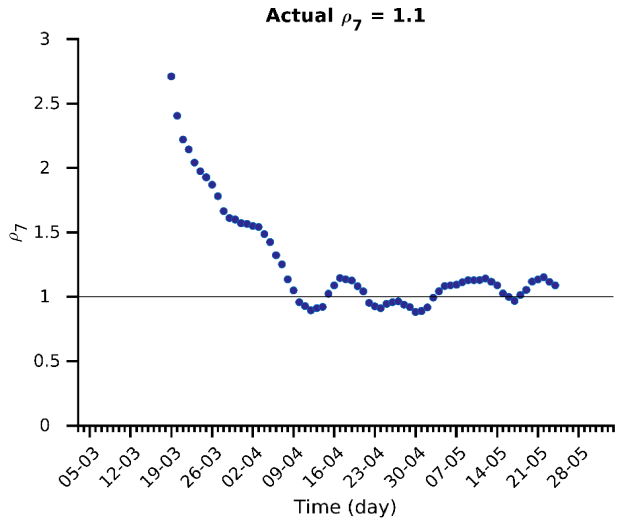


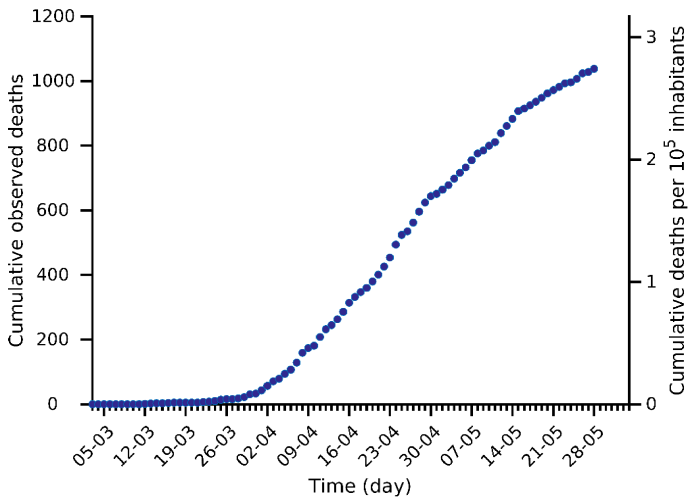


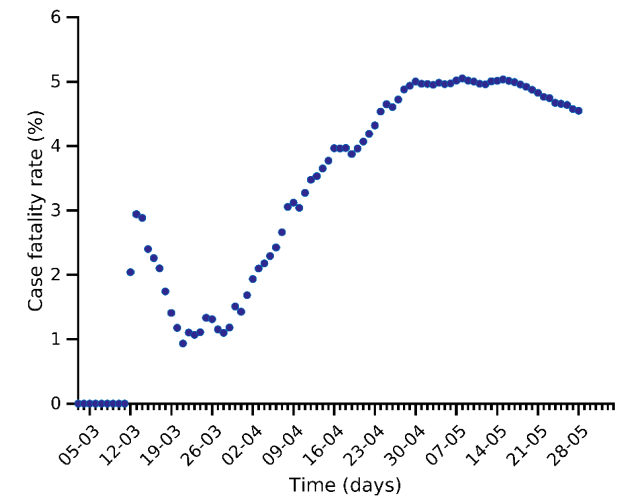




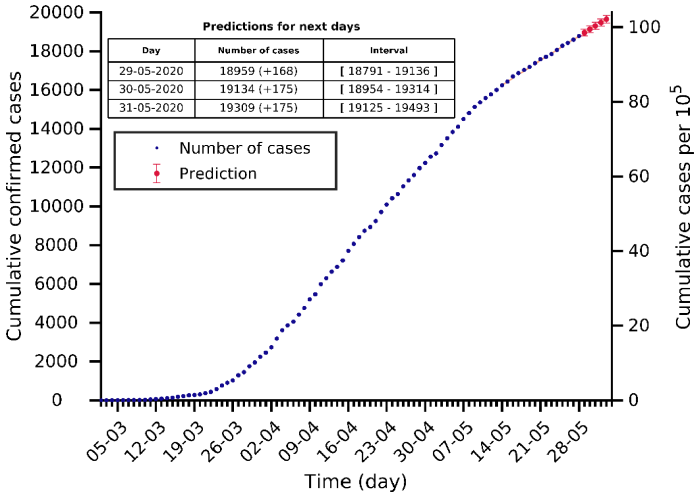


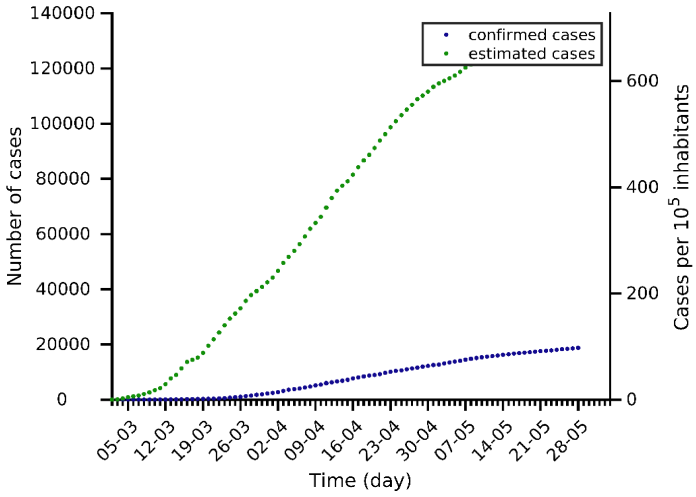


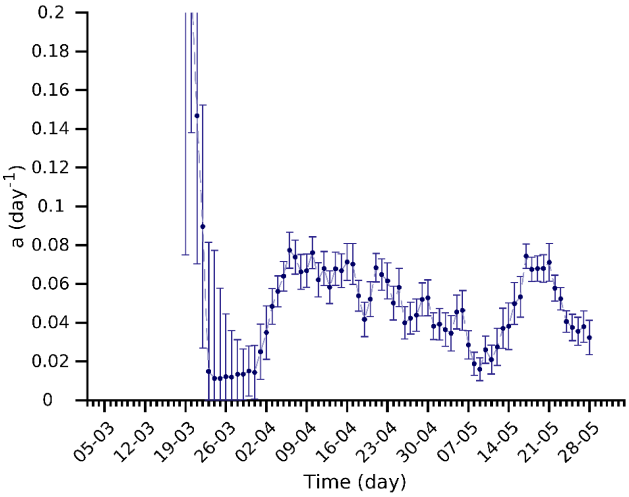


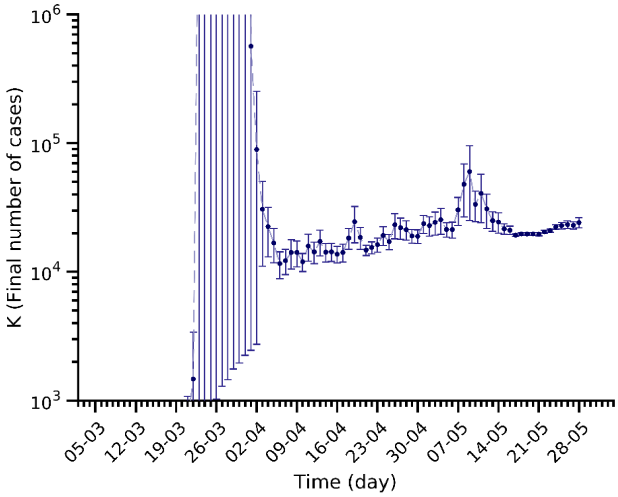


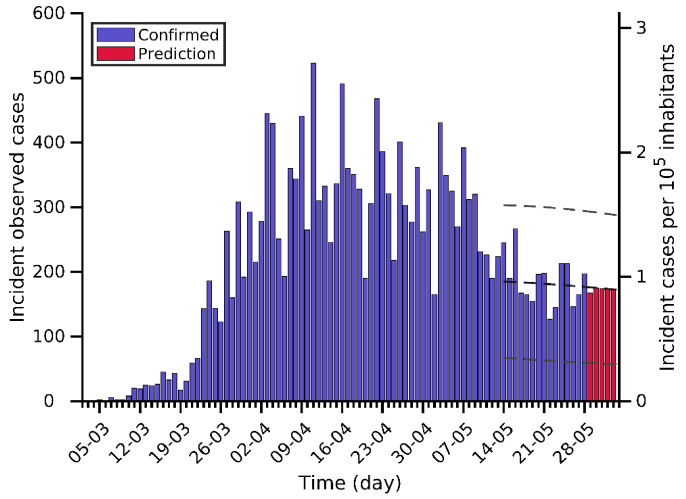
Romania 28-05-2020. Population: 19.2M. Current cumulated incidence: 98/10^5

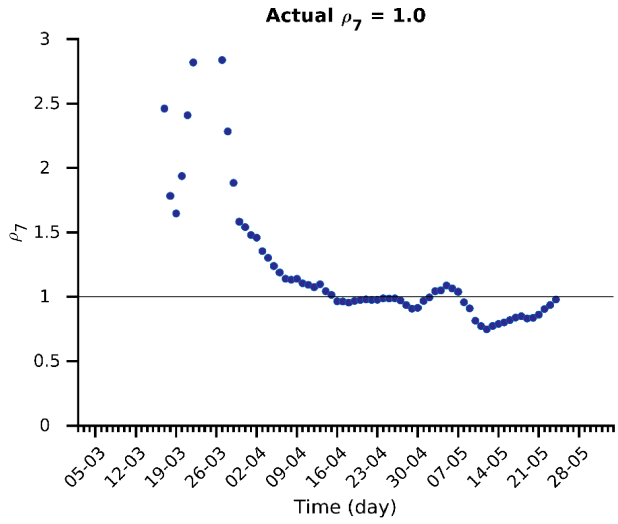


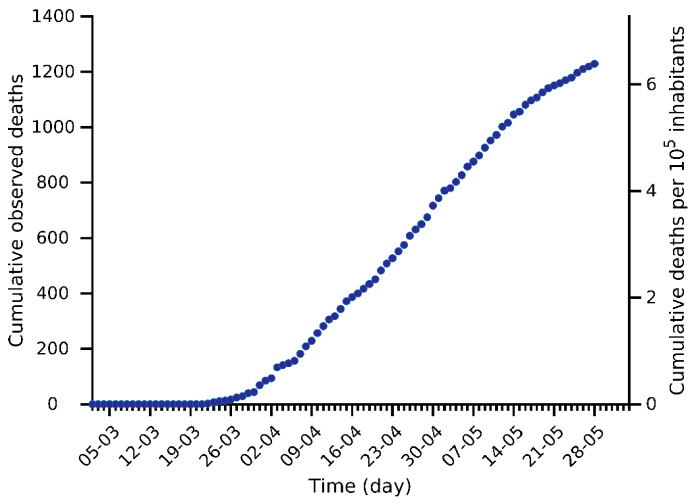


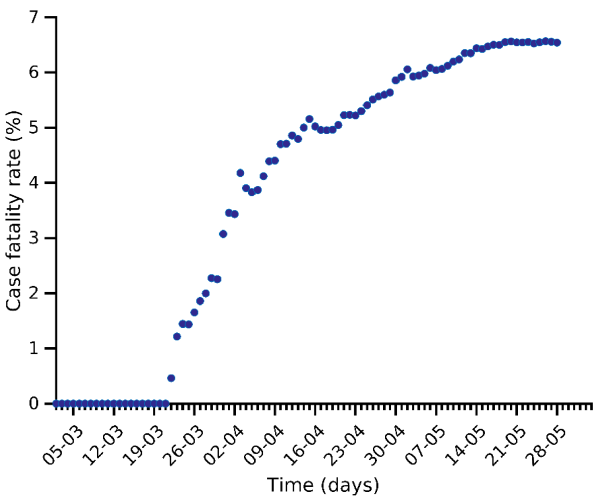




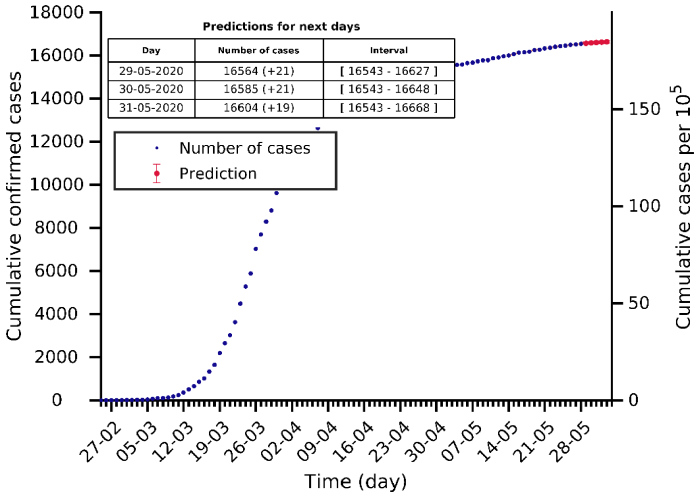


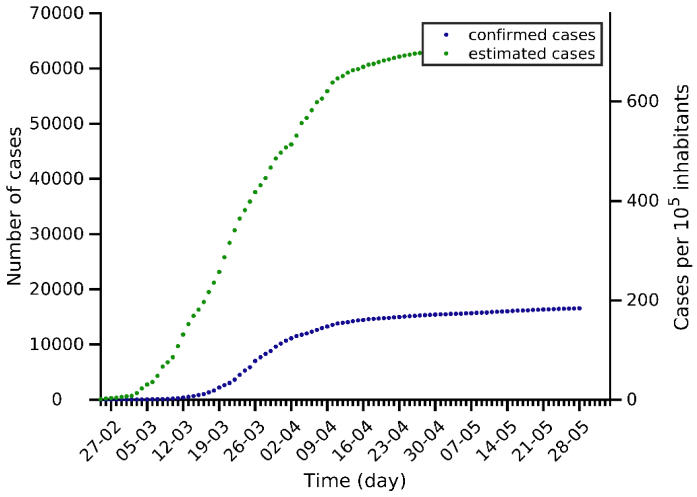


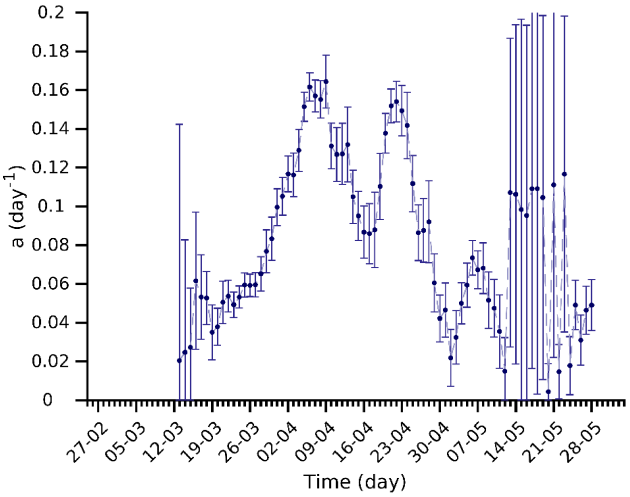


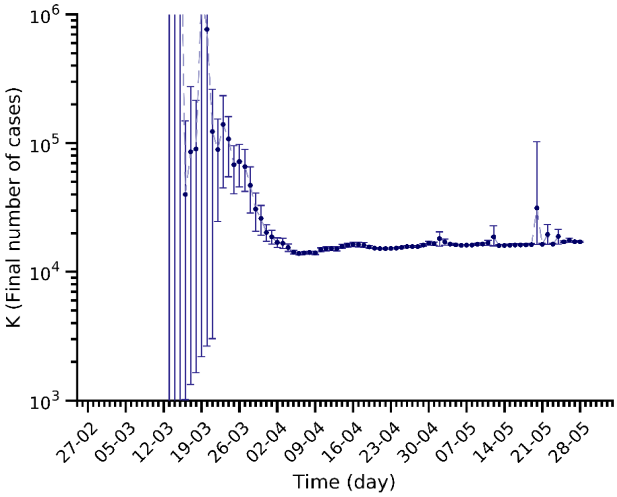


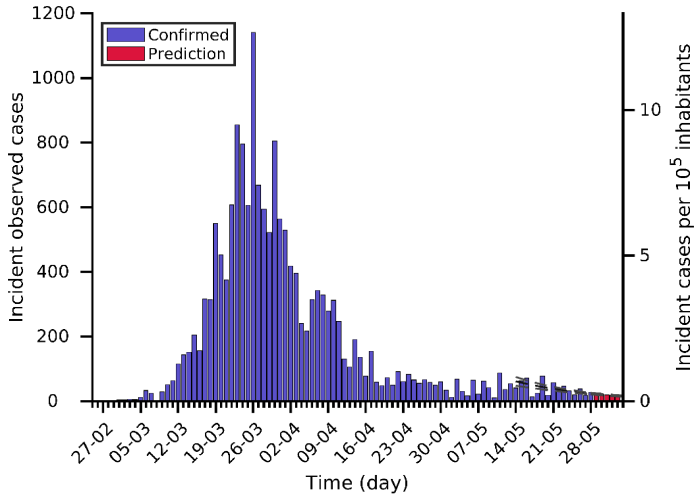
Austria 28-05-2020. Population: 9.0M. Current cumulated incidence: 184/10^5

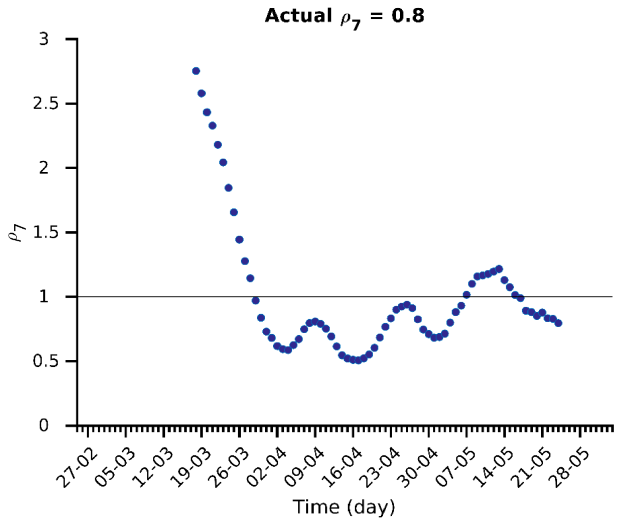


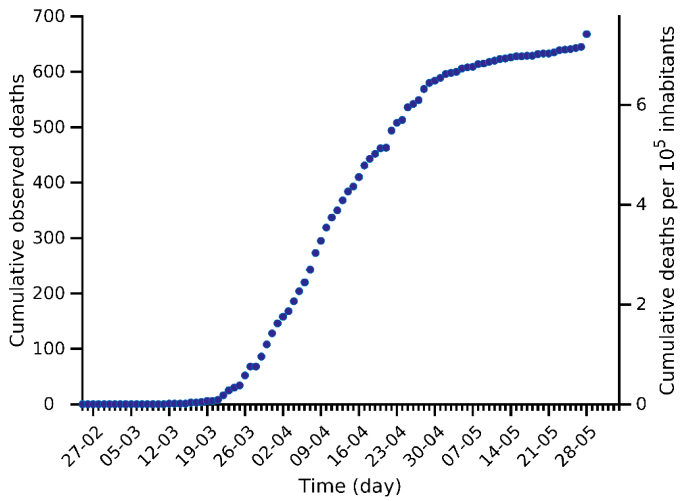


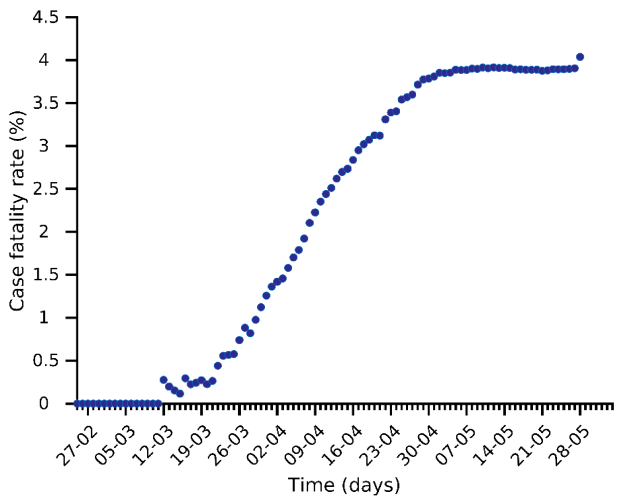




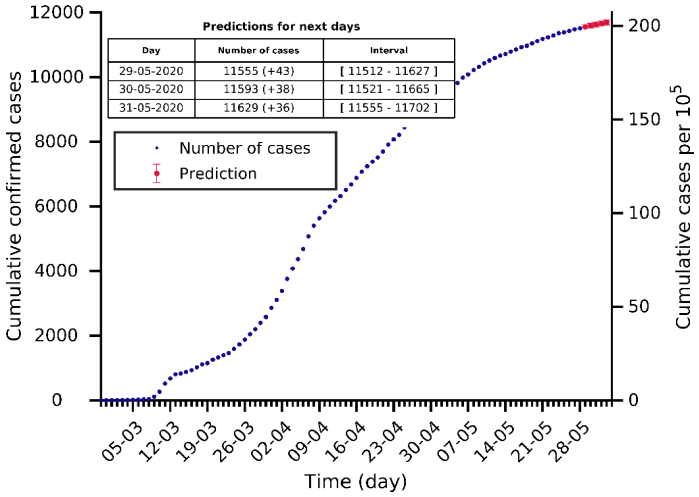


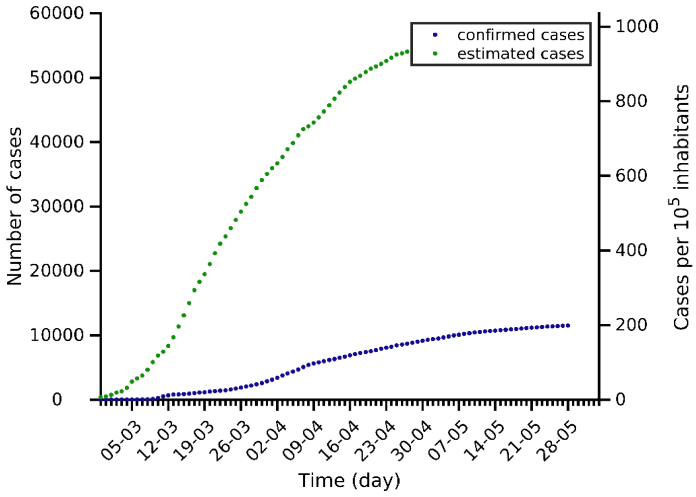


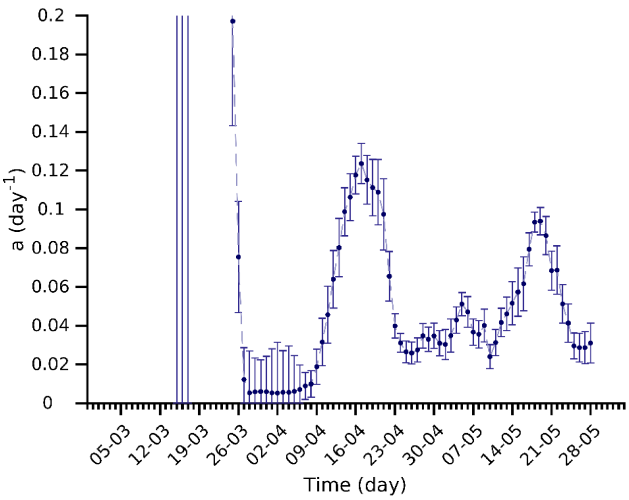


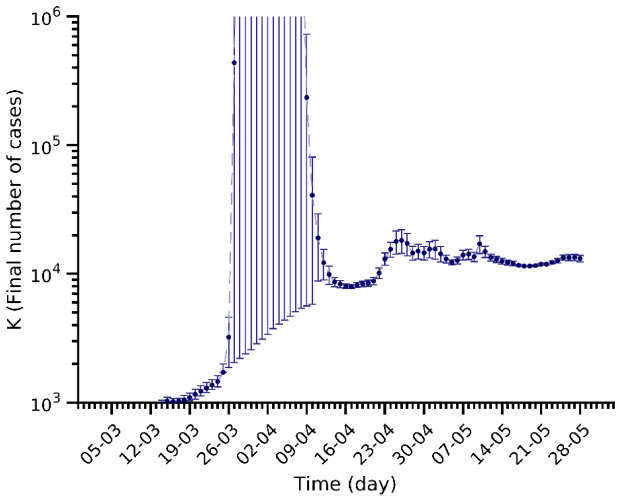


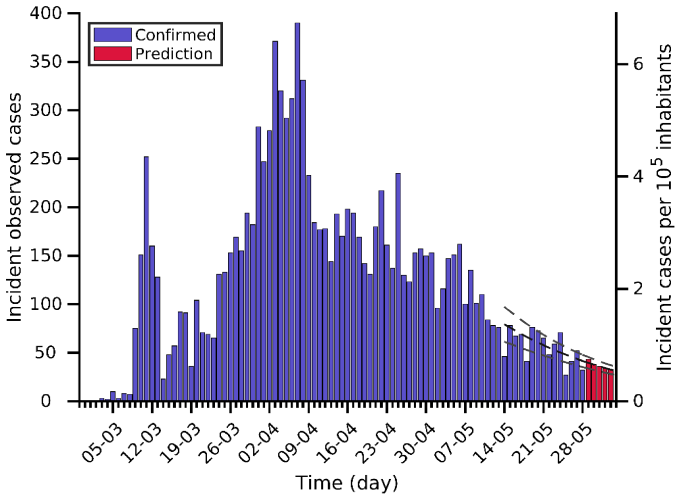
Denmark 28-05-2020. Population: 5.8M. Current cumulated incidence: 199/10^5

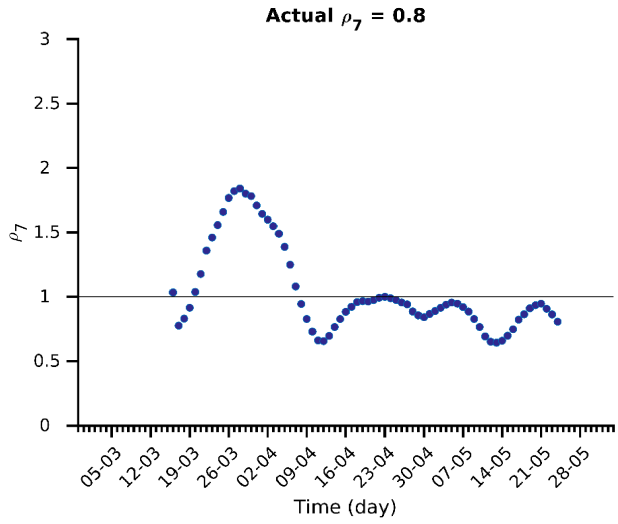


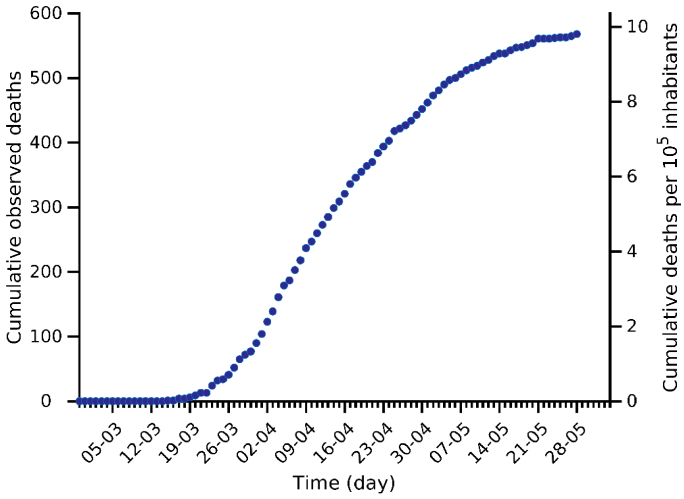


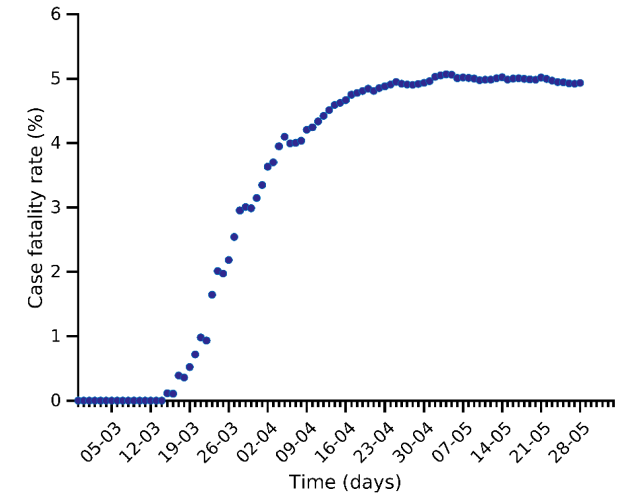




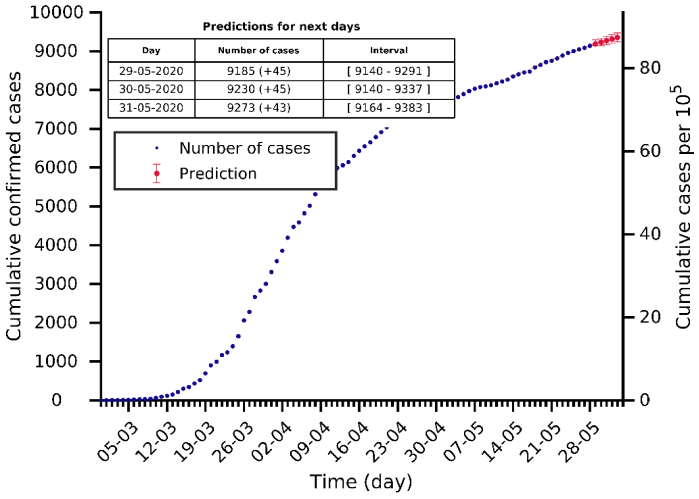


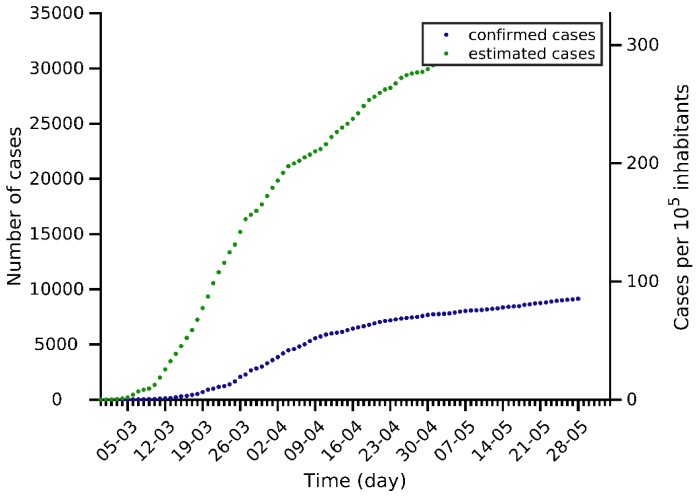


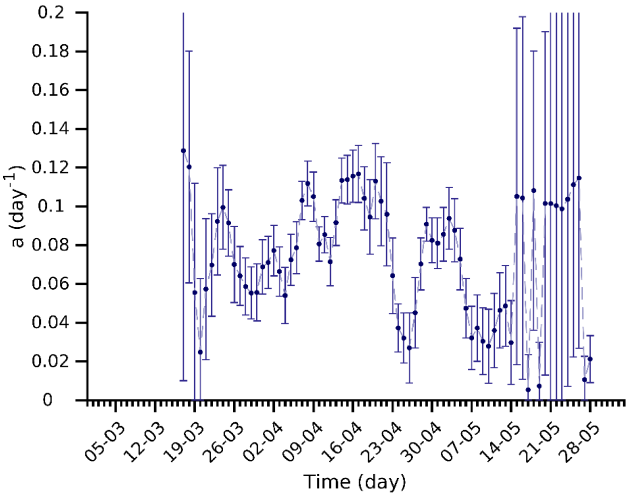


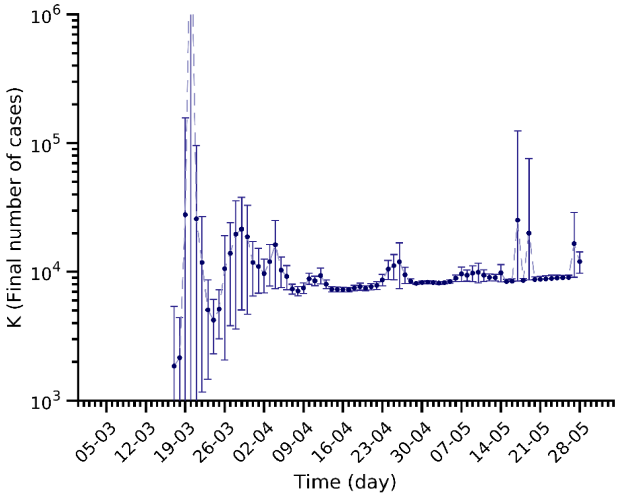


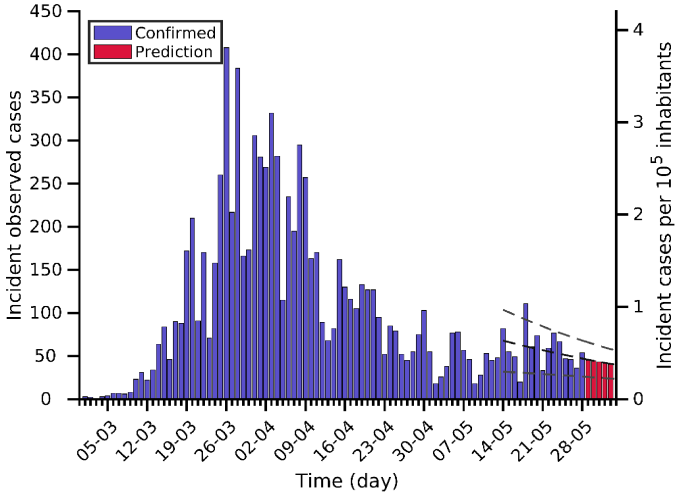
Czech Rep 28-05-2020. Population: 10.7M. Current cumulated incidence: 85/10^5

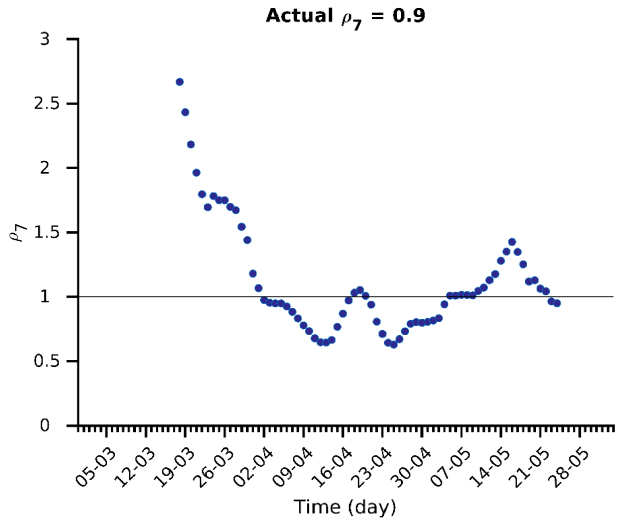


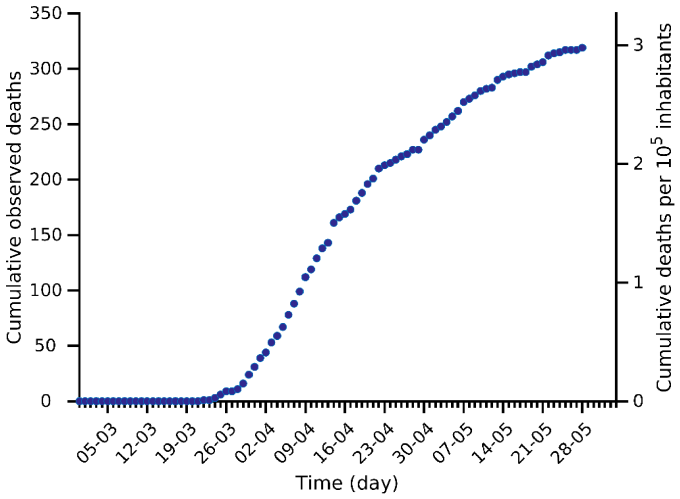


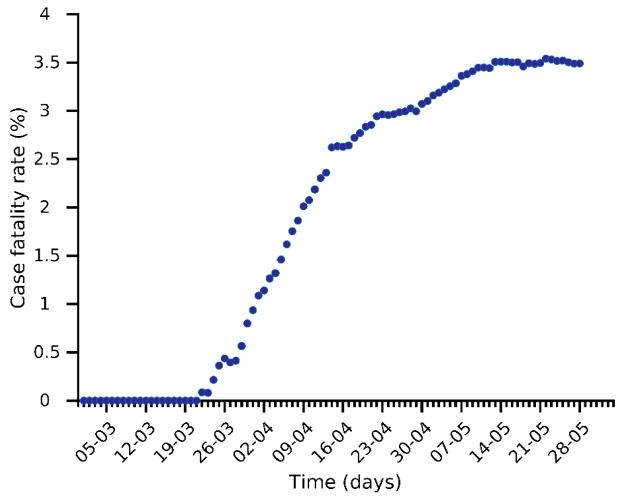




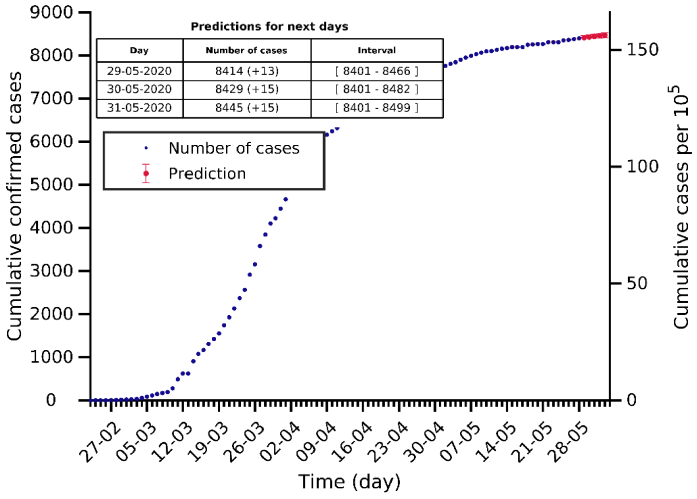


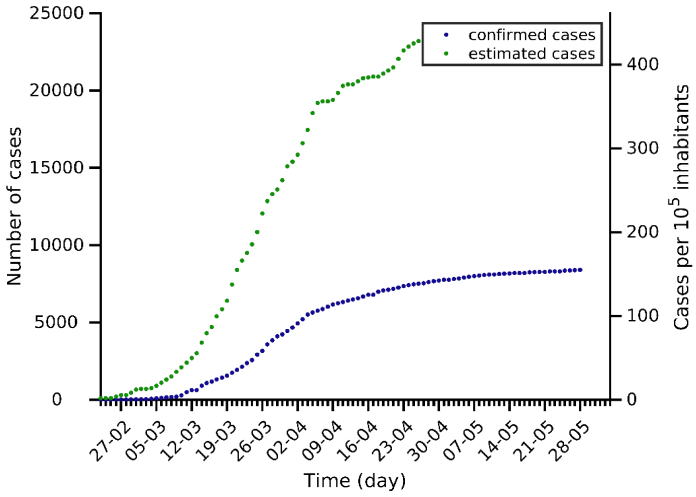


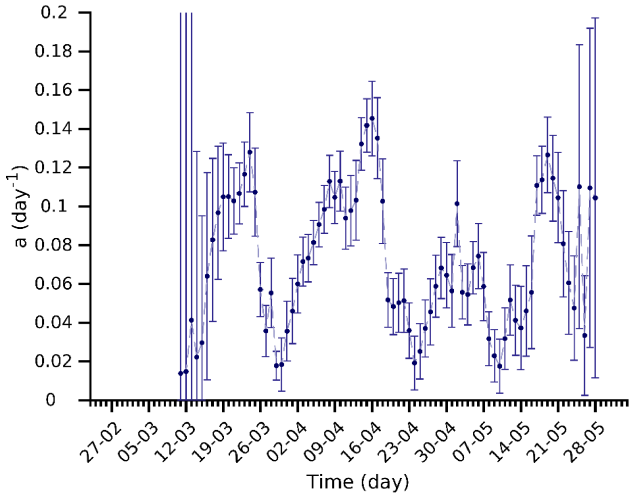


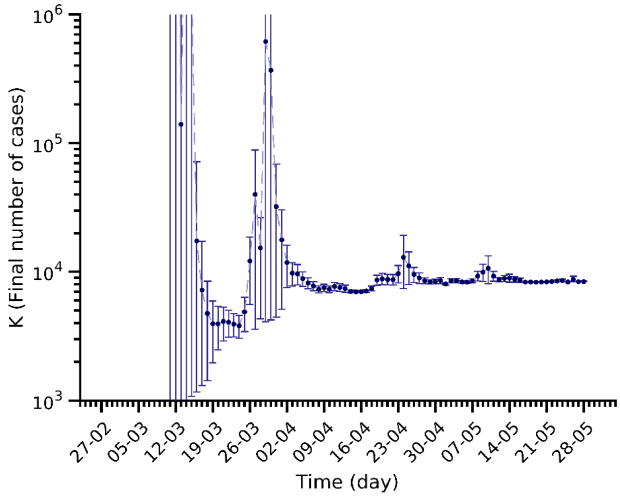


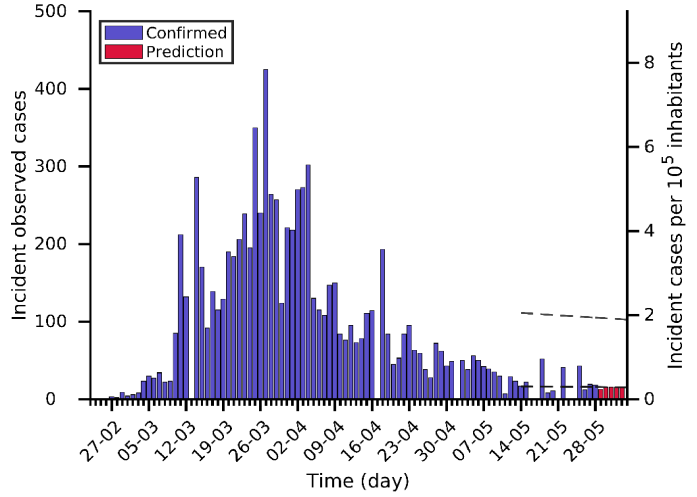
Norway 28-05-2020. Population: 5.4M. Current cumulated incidence: 155/10⁵

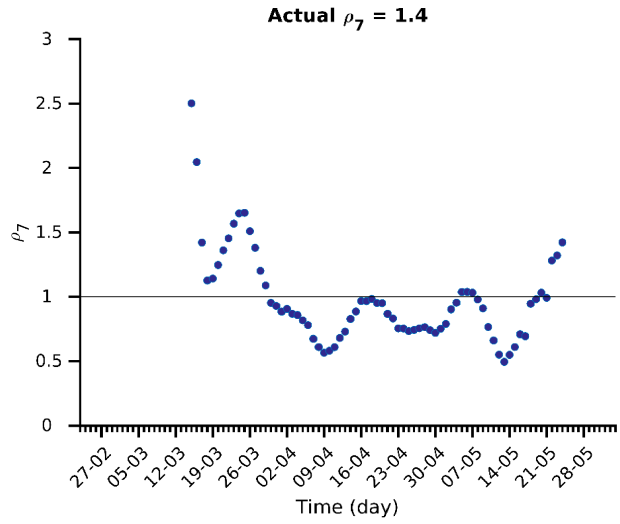


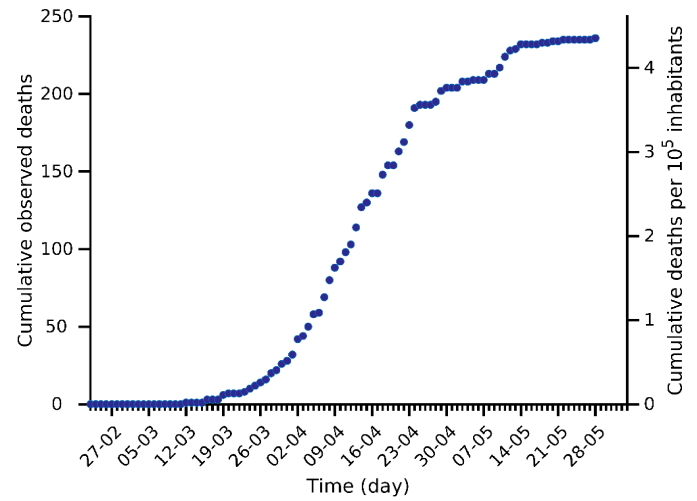


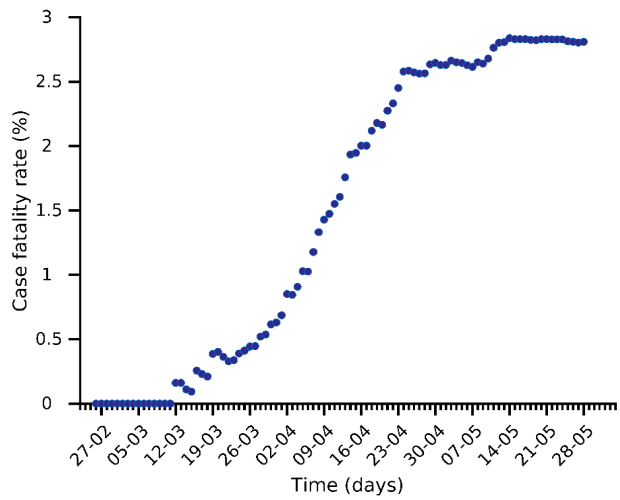




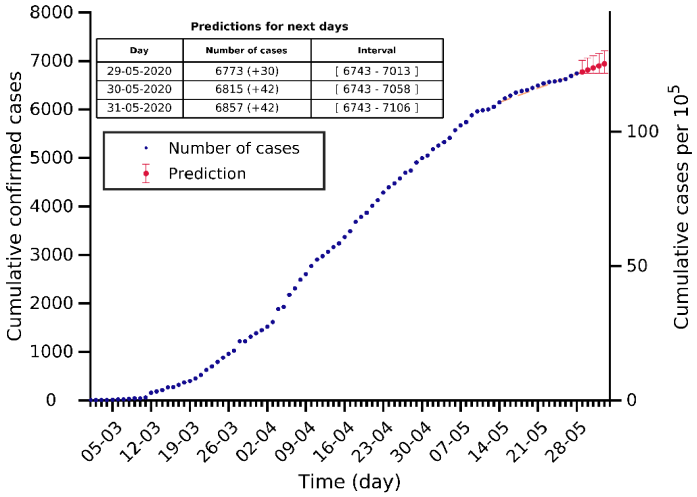


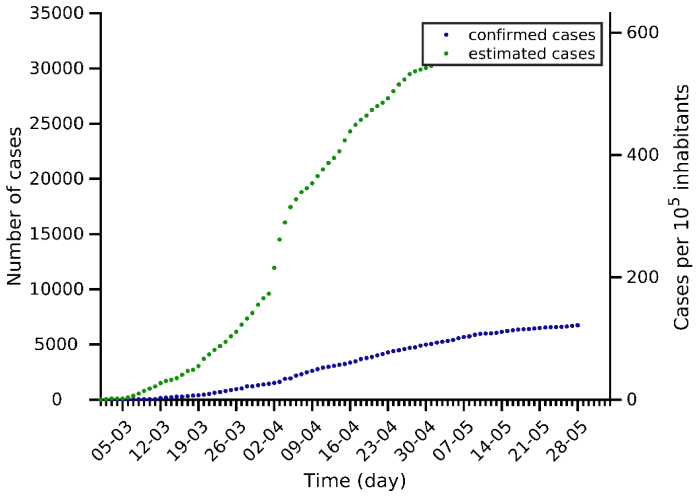


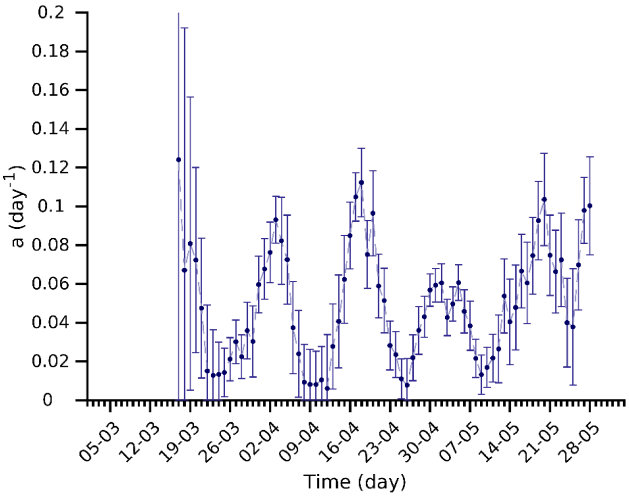


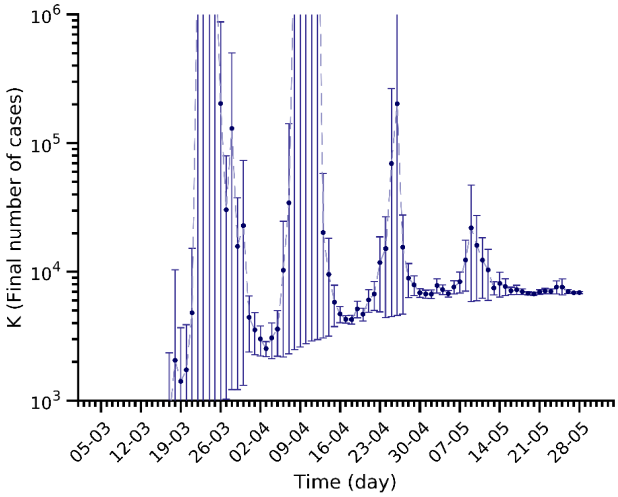


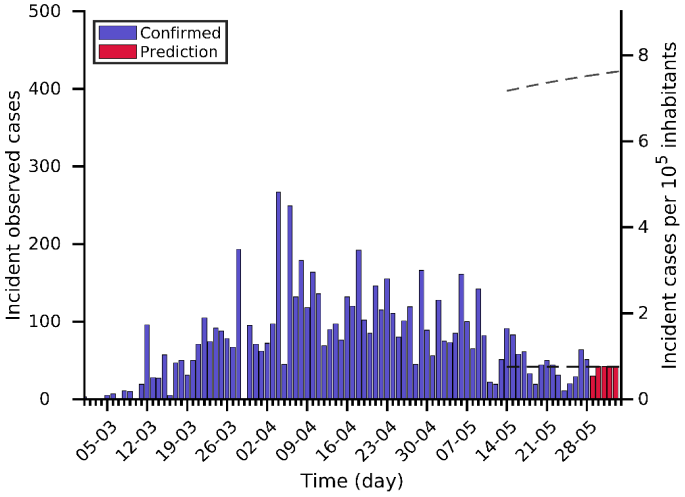
Finland 28-05-2020. Population: 5.5M. Current cumulated incidence: 122/10⁵

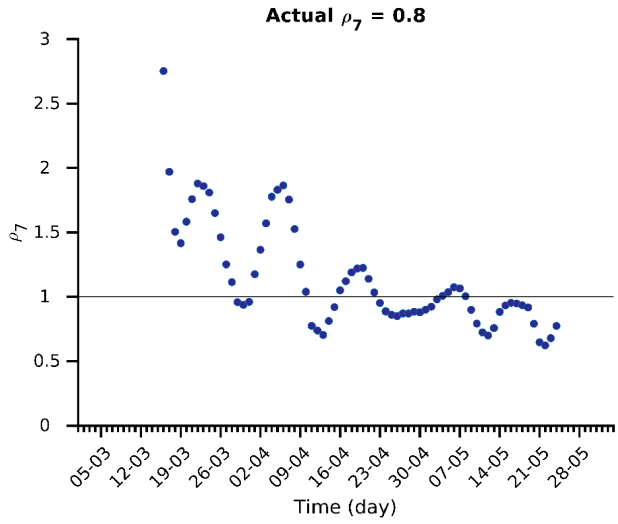


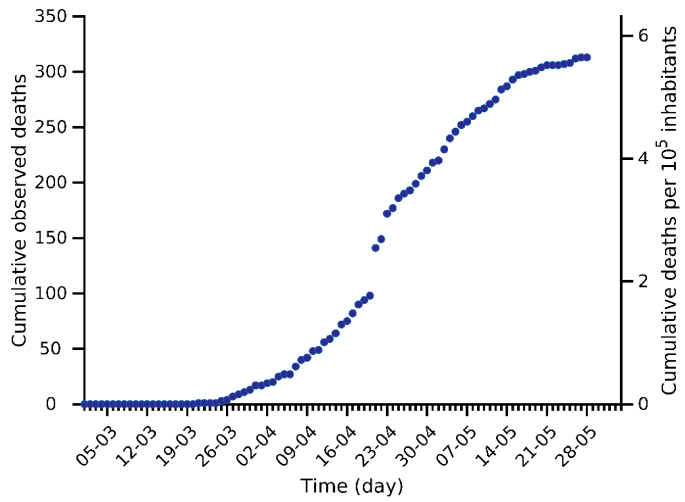


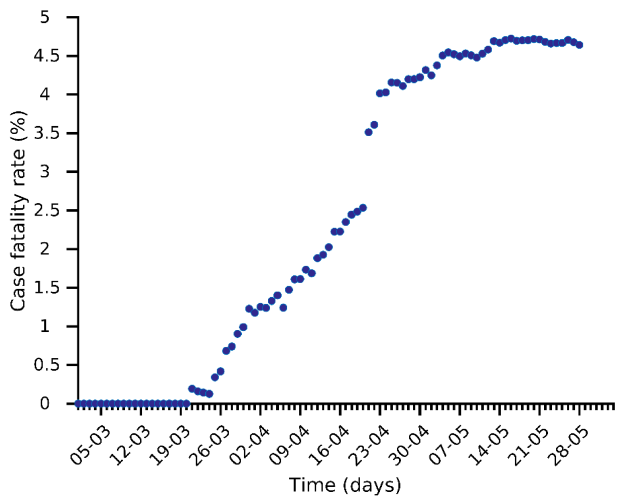




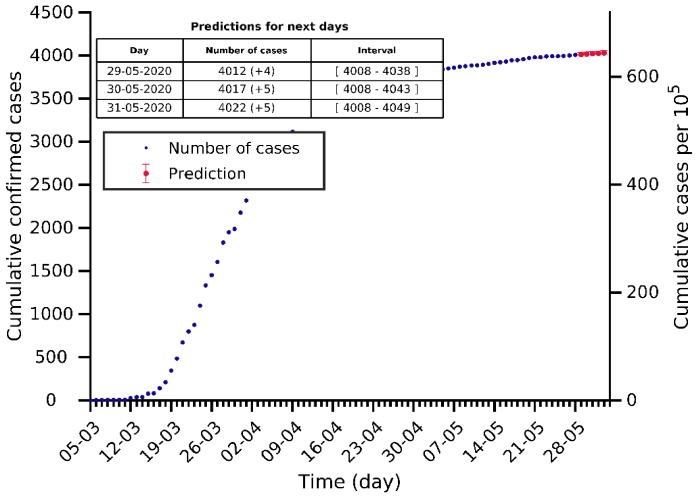


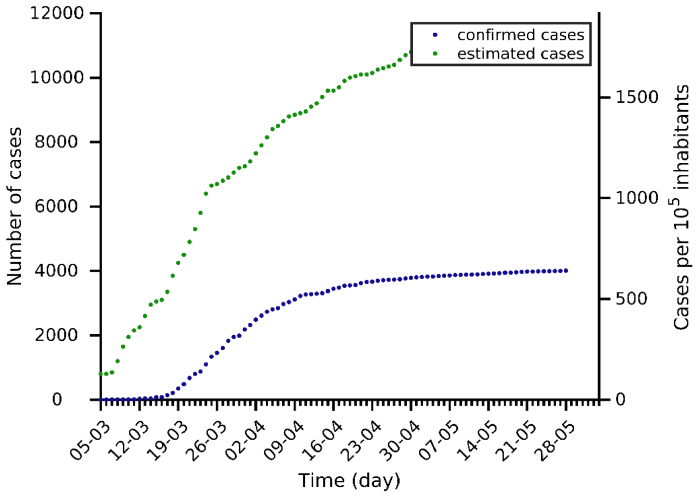


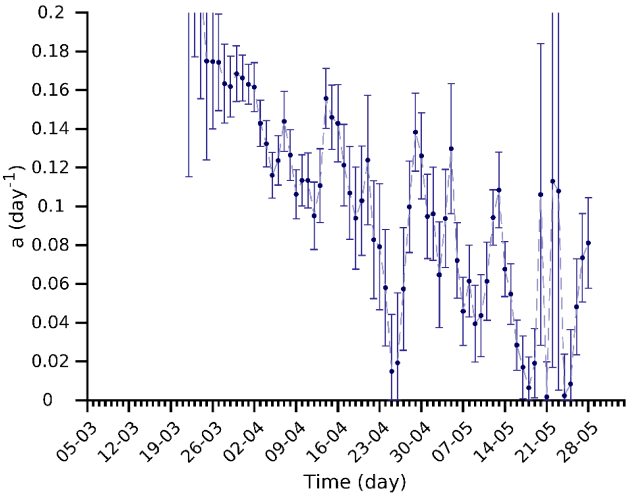


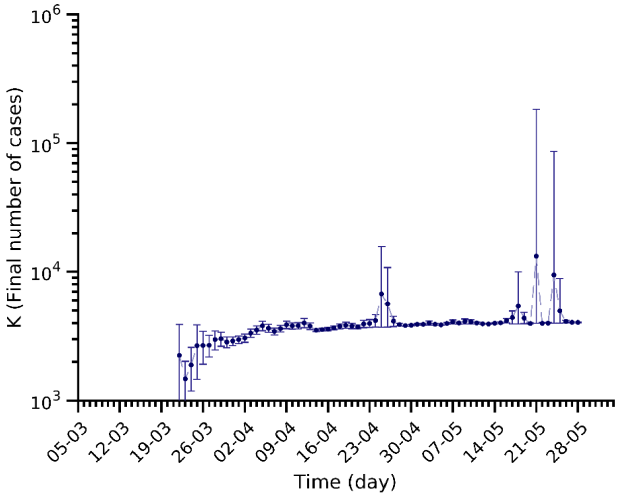


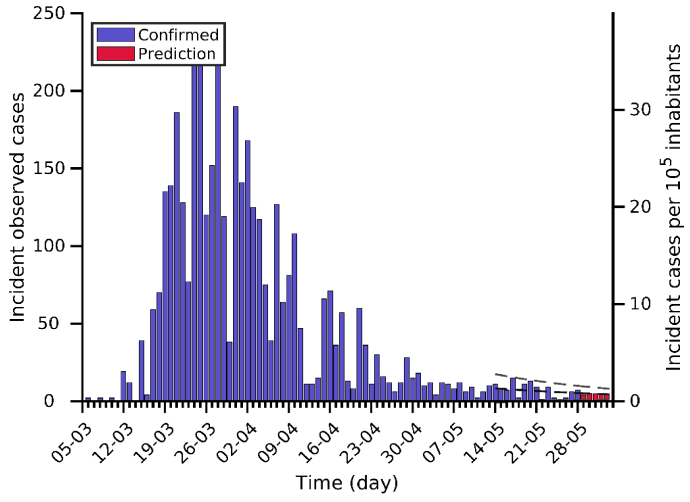
Luxembourg 28-05-2020. Population: 0.6M. Current cumulated incidence: 640/10^5

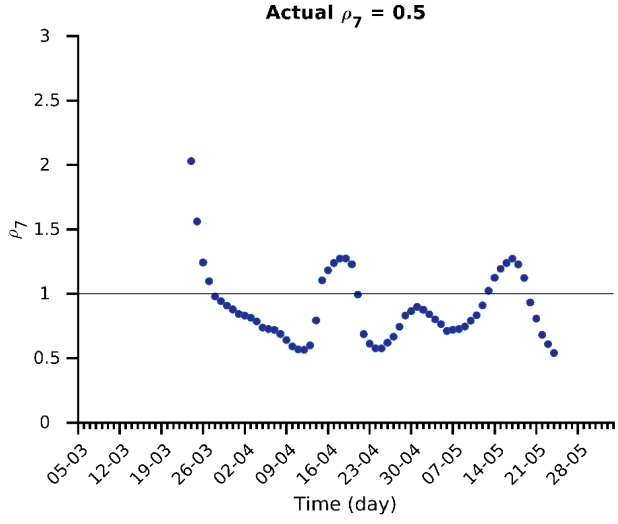


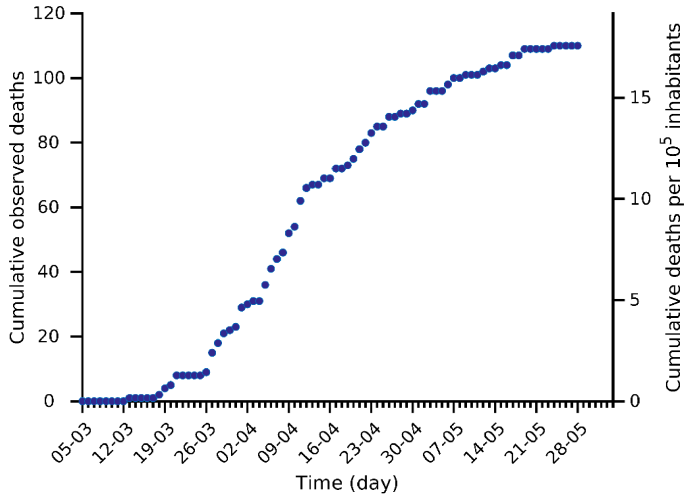


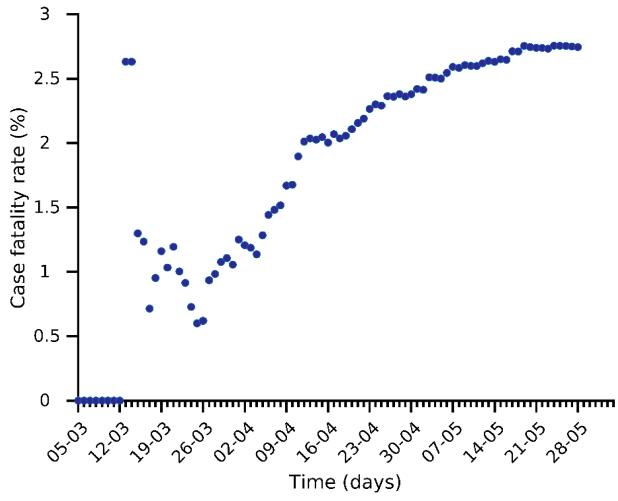




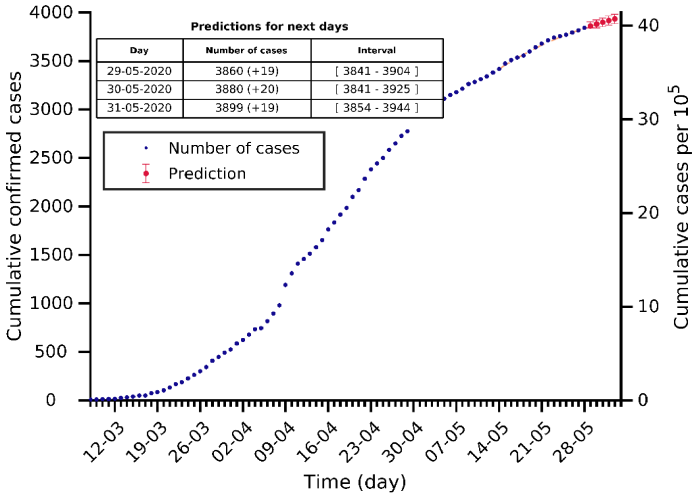


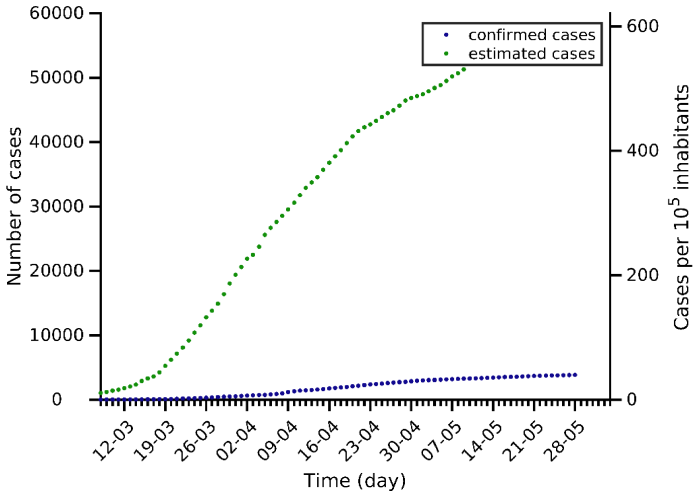


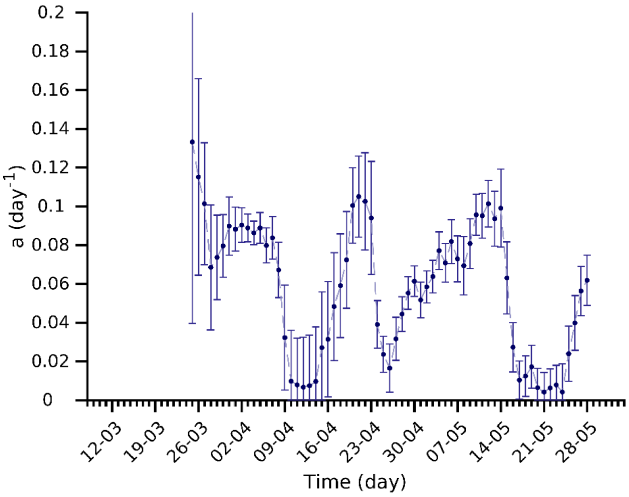


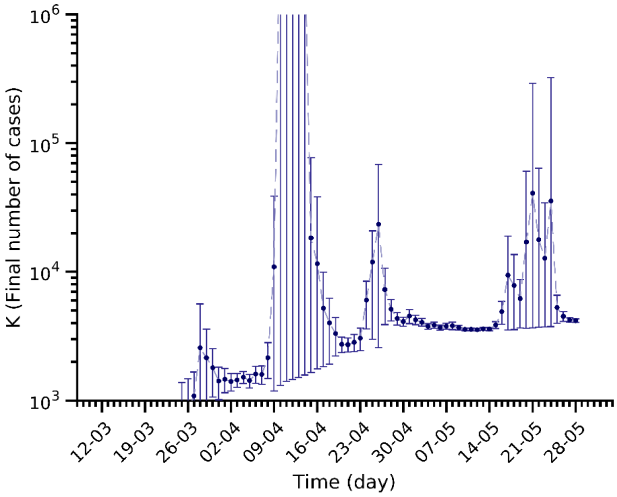


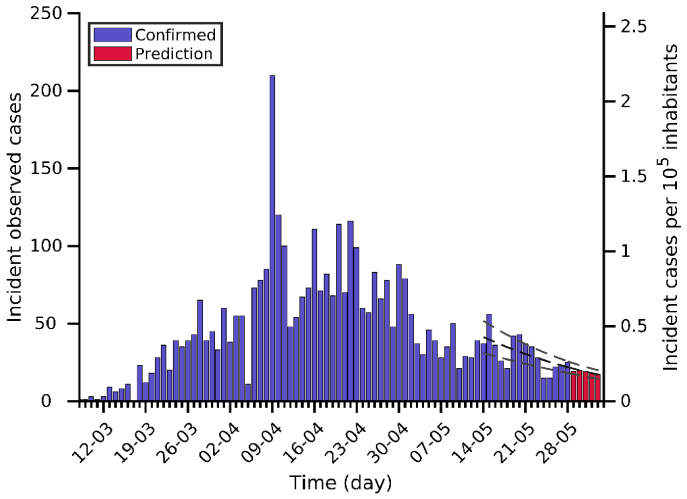
Hungary 28-05-2020. Population: 9.7M. Current cumulated incidence: $40/10^5$

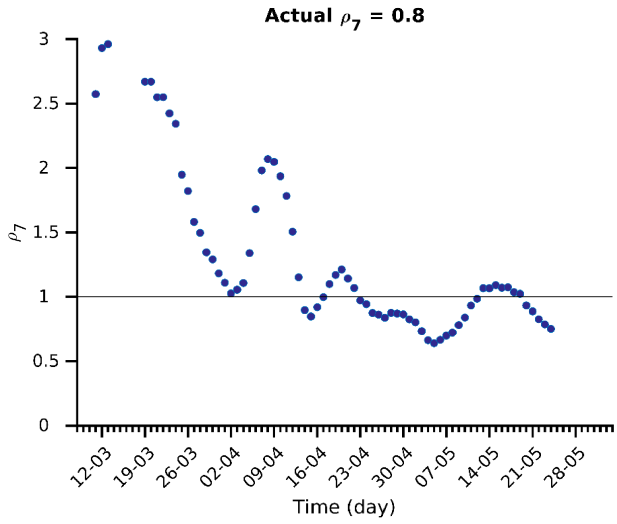


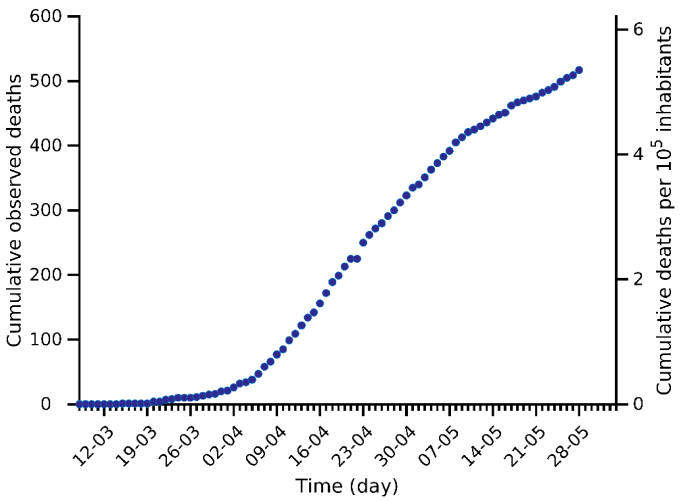


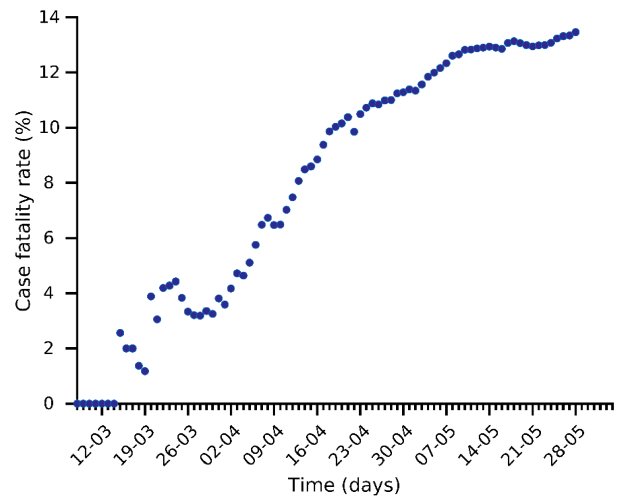




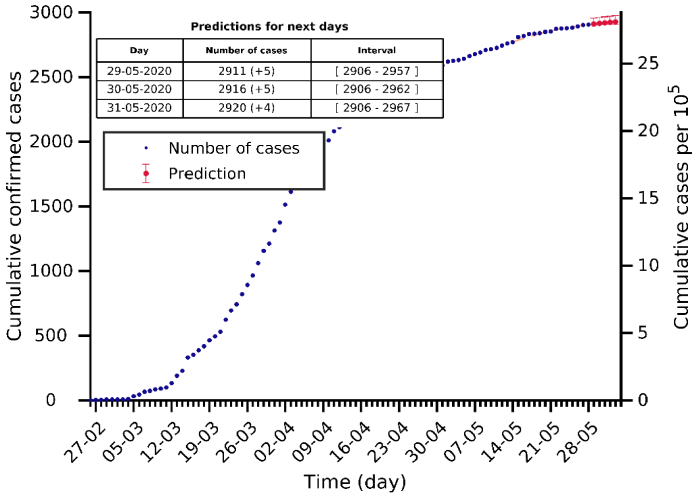


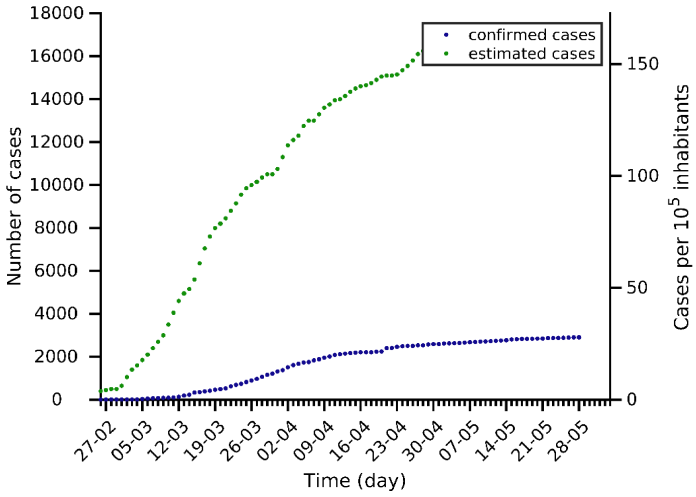


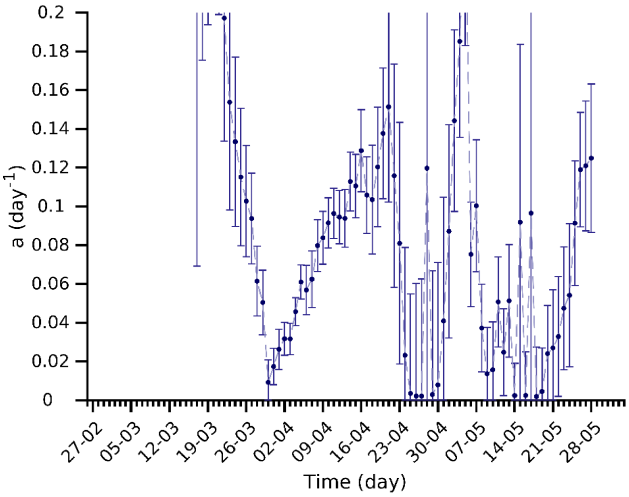


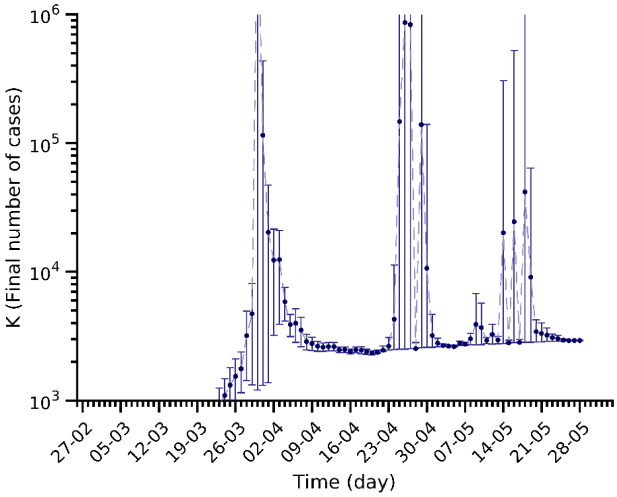


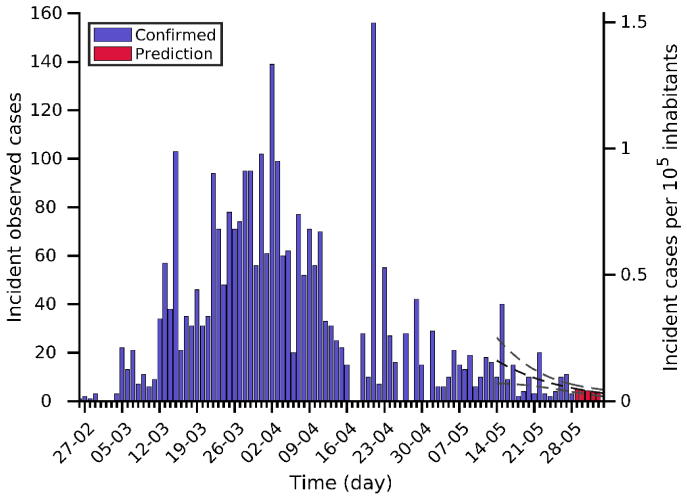
Greece 28-05-2020. Population: 10.4M. Current cumulated incidence: $28/10^5$

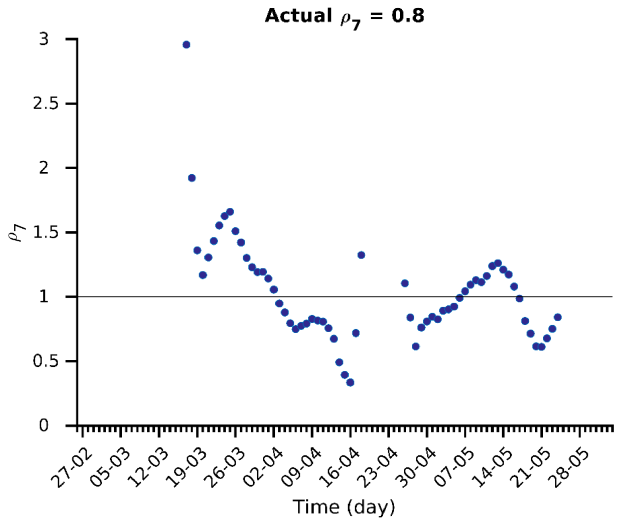


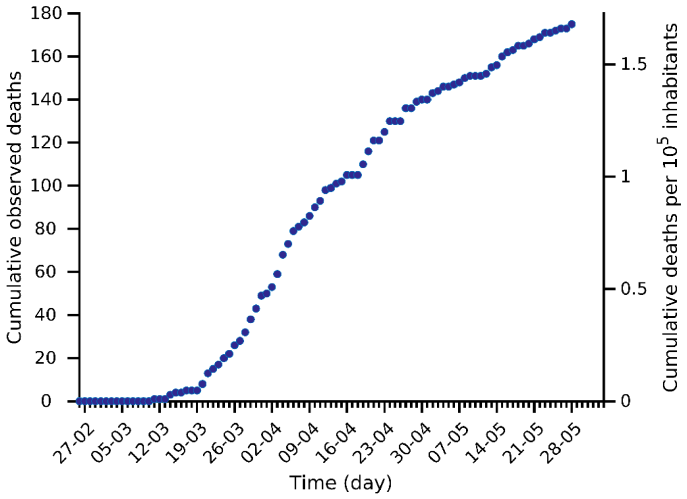


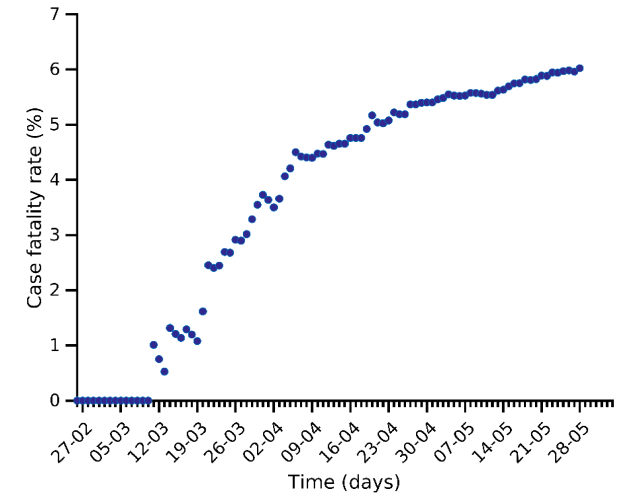




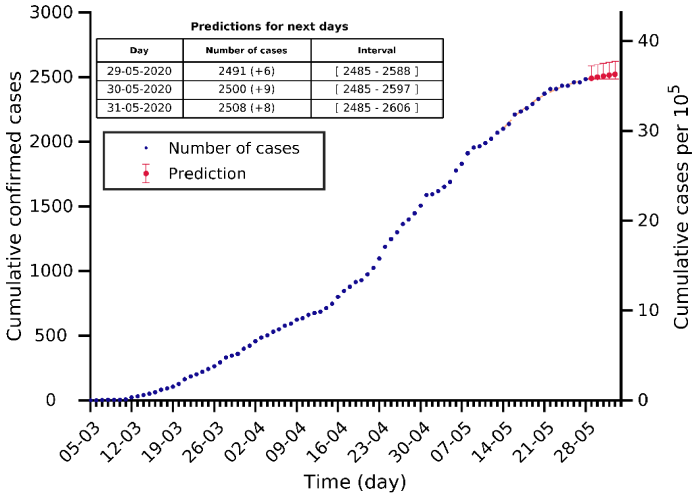


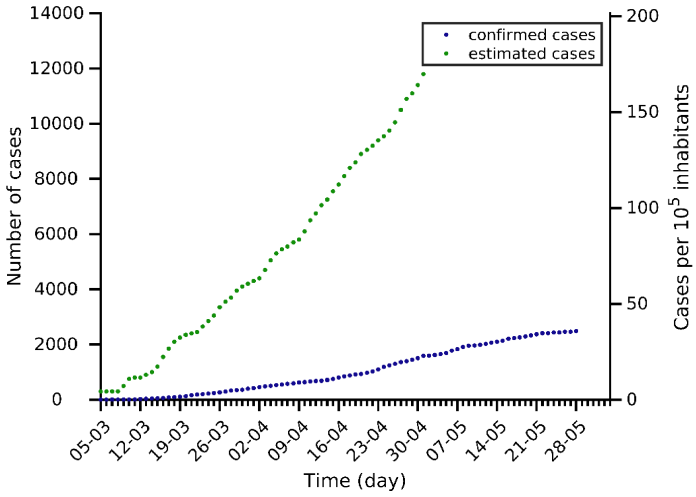


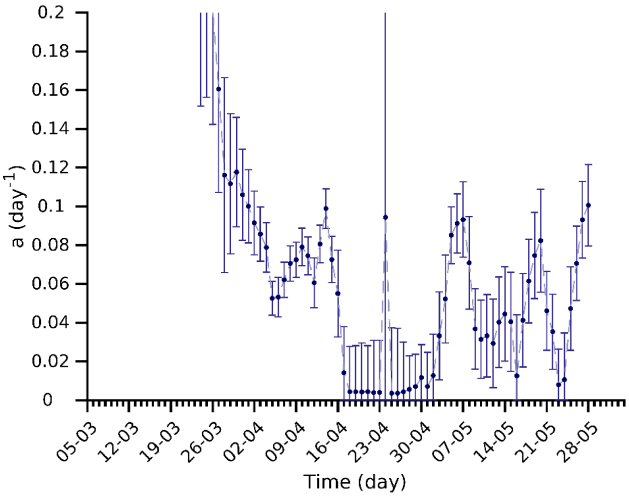


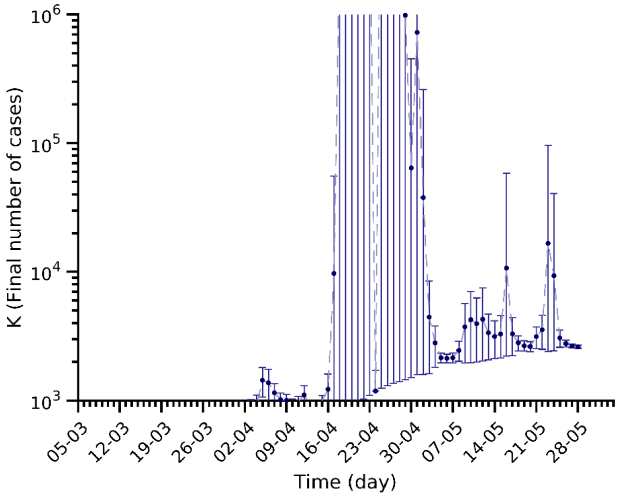


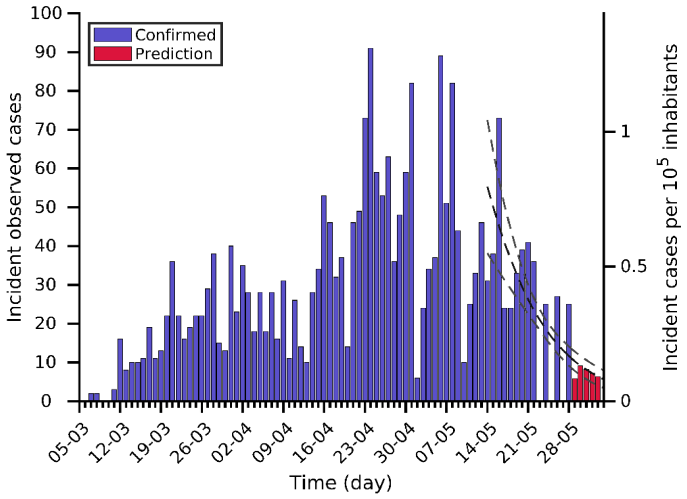
Bulgaria 28-05-2020. Population: 6.9M. Current cumulated incidence: $36/10^5$

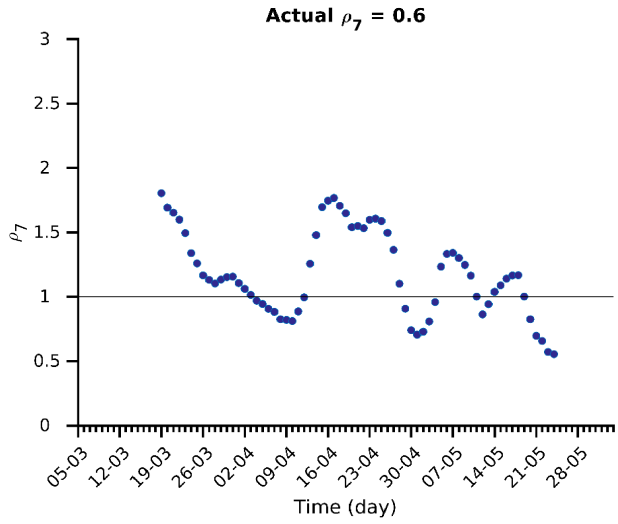


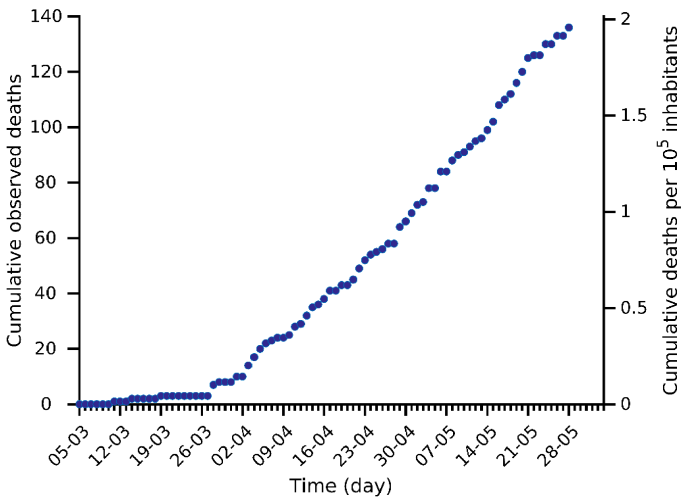


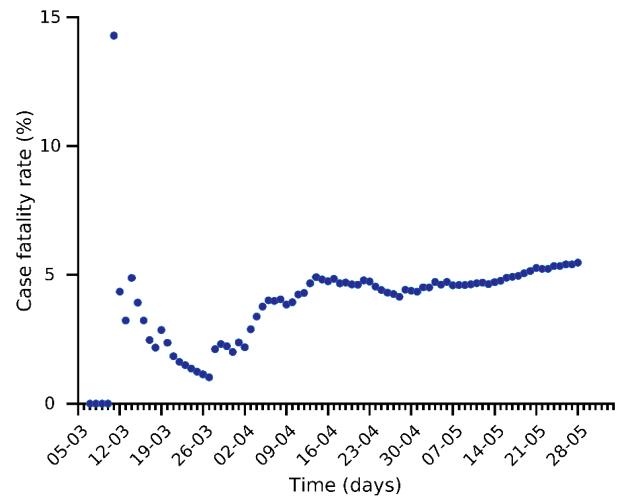




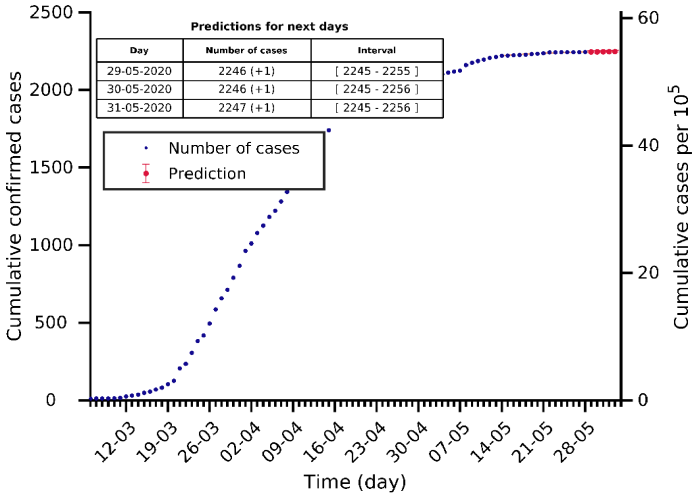


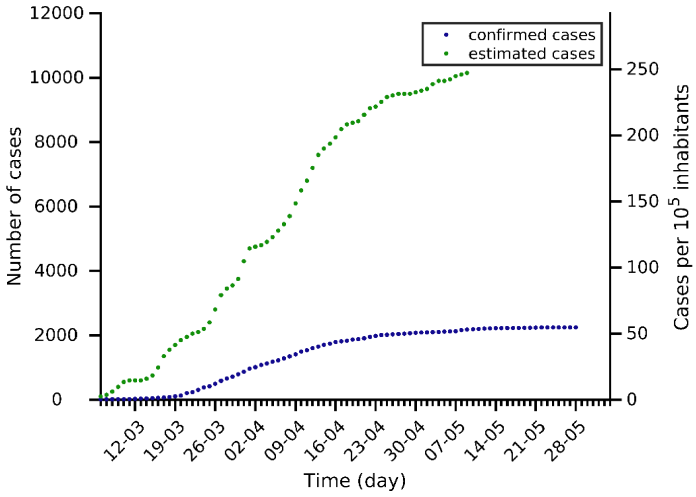


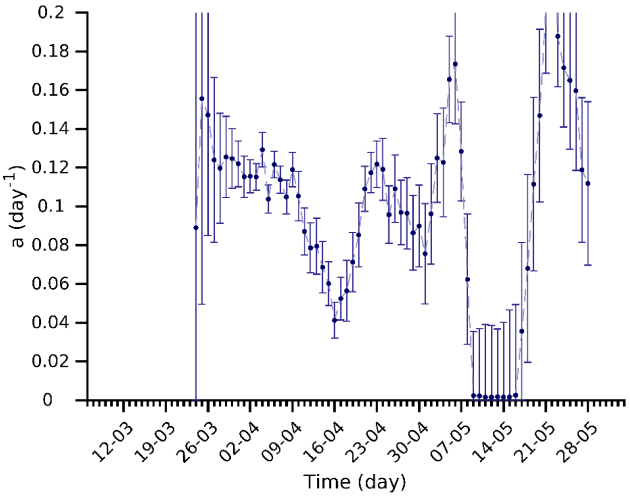


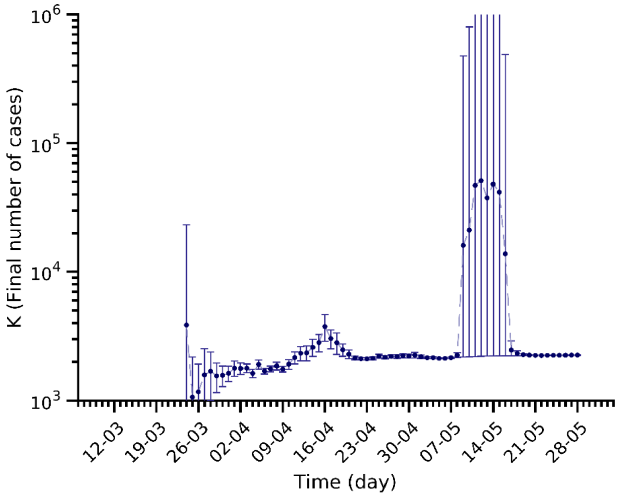


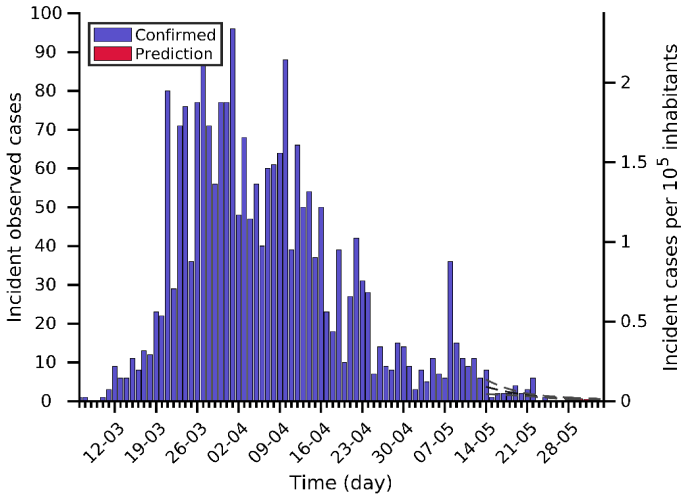
Croatia 28-05-2020. Population: 4.1M. Current cumulated incidence: $55/10^5$

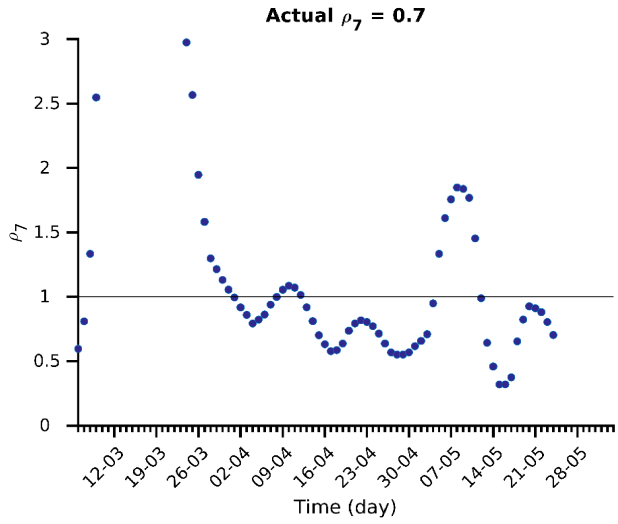


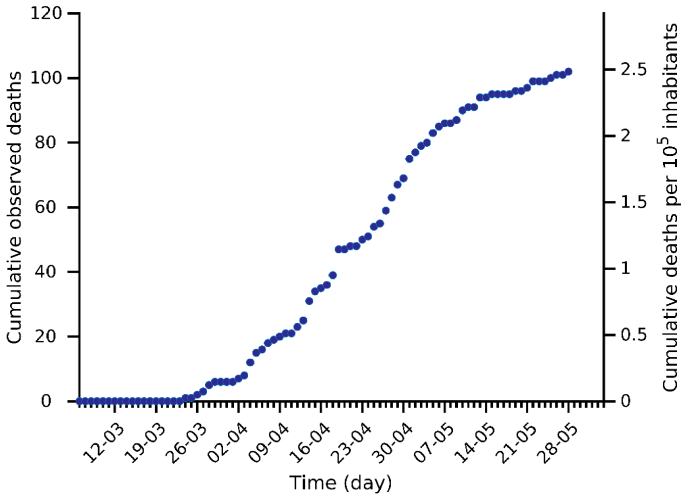


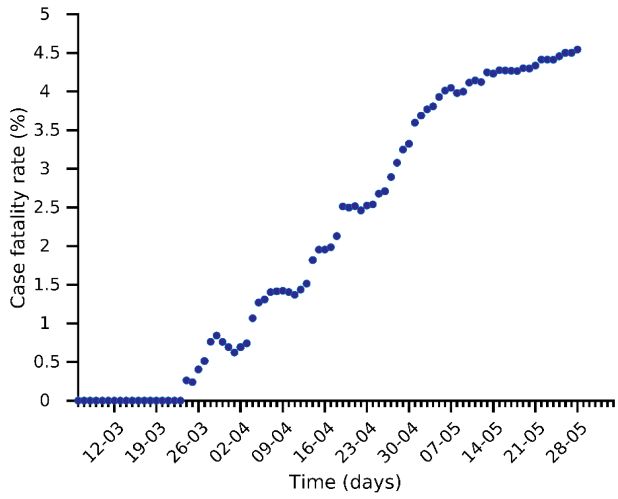




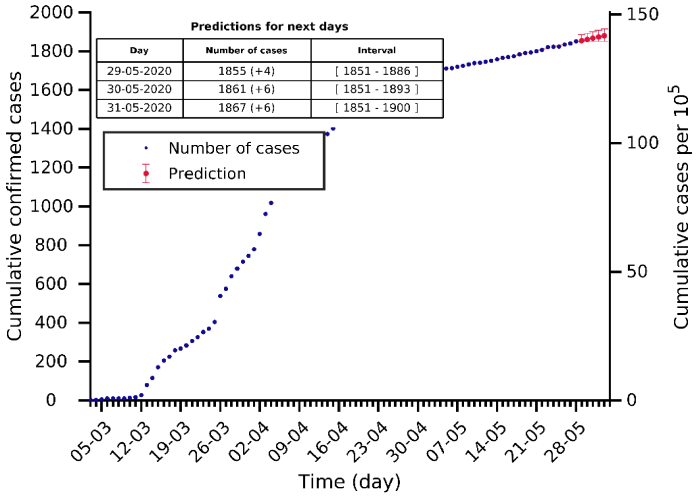


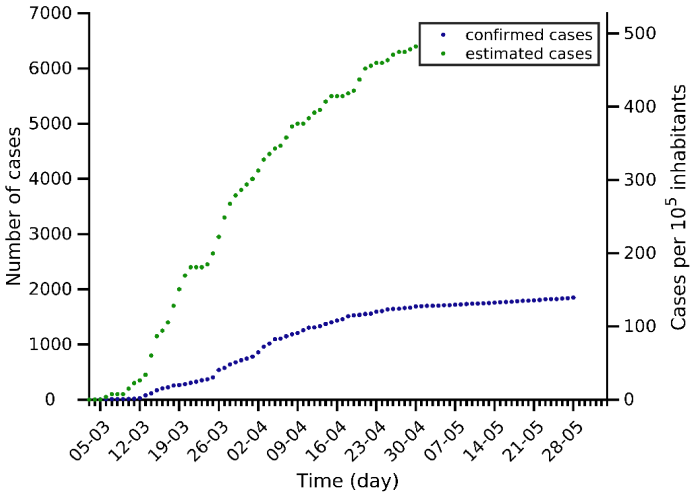


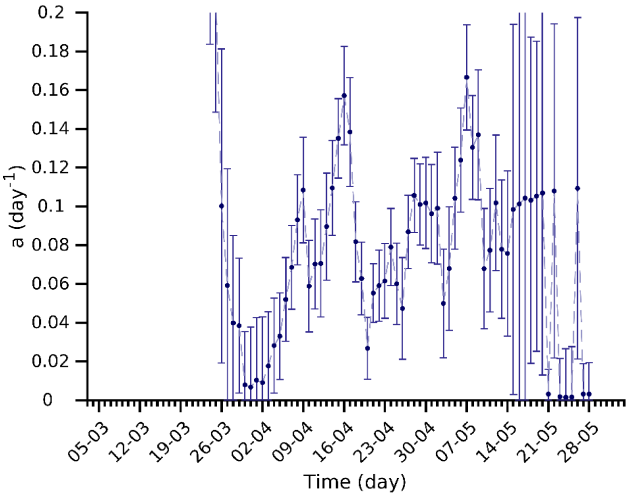


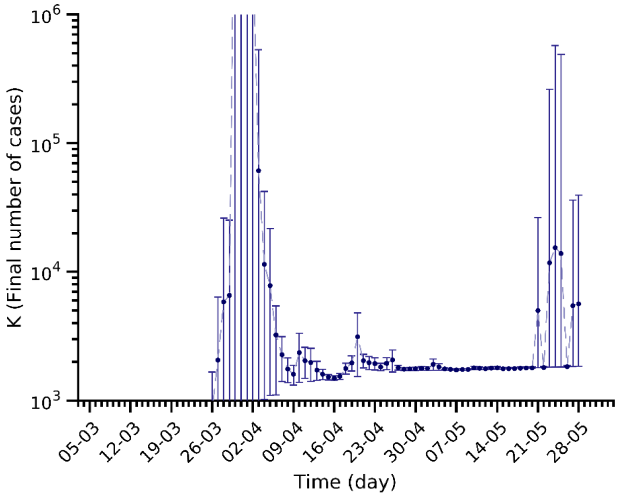


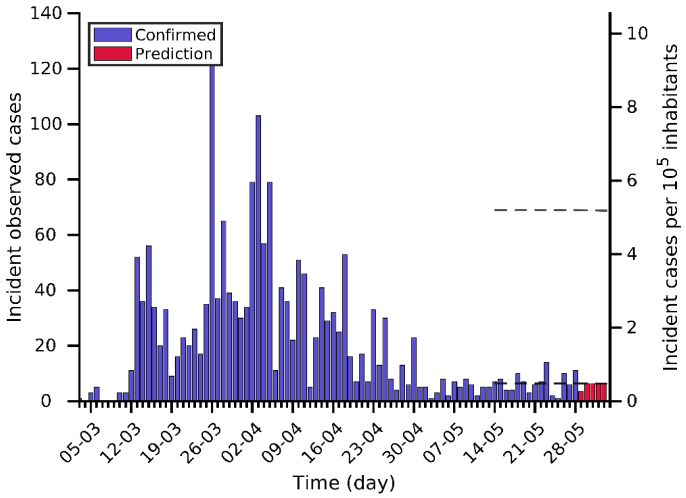
Estonia 28-05-2020. Population: 1.3M. Current cumulated incidence: 140/10^5

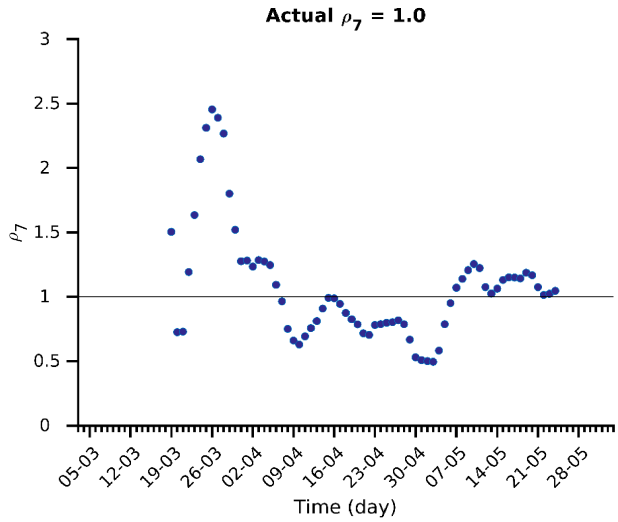


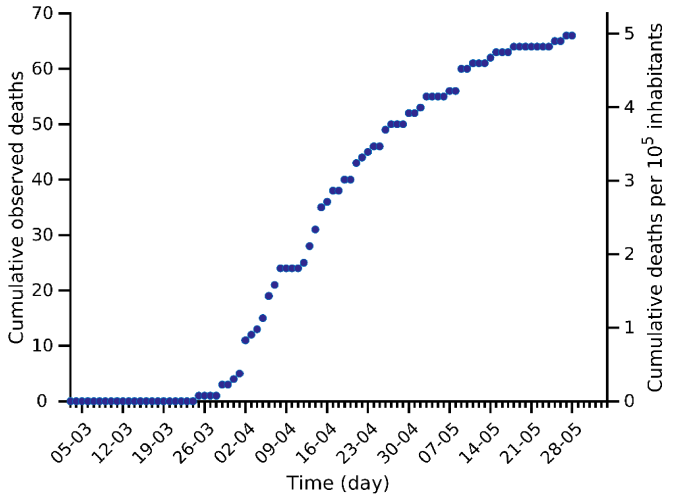


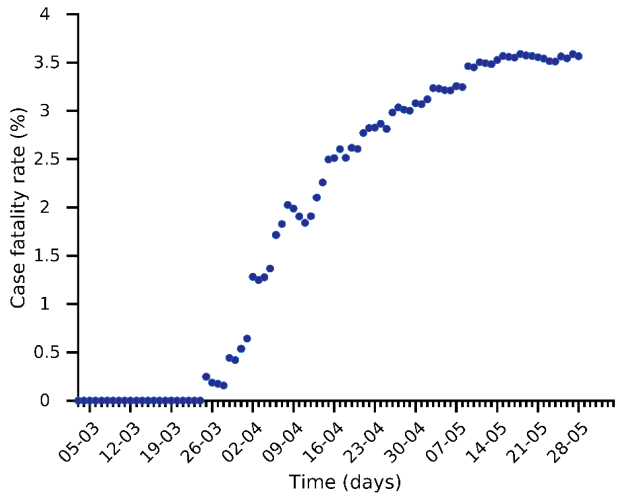




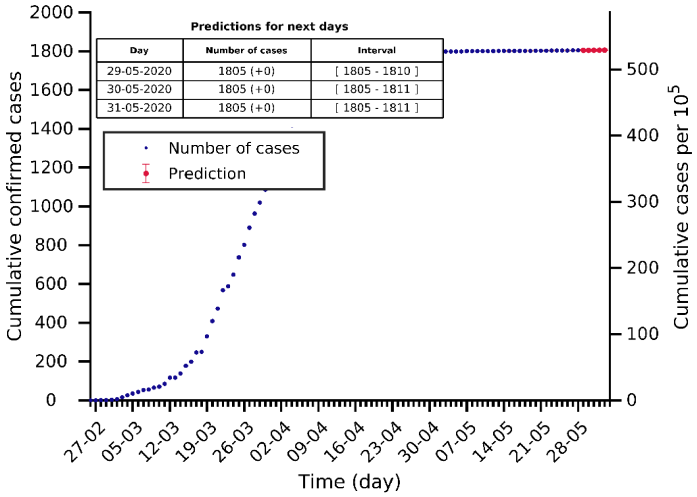


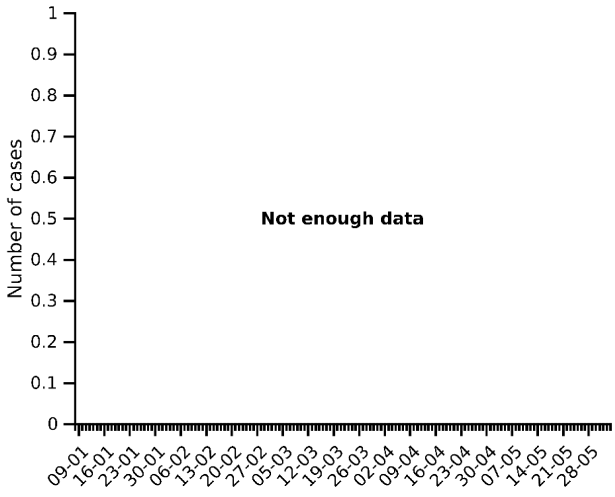


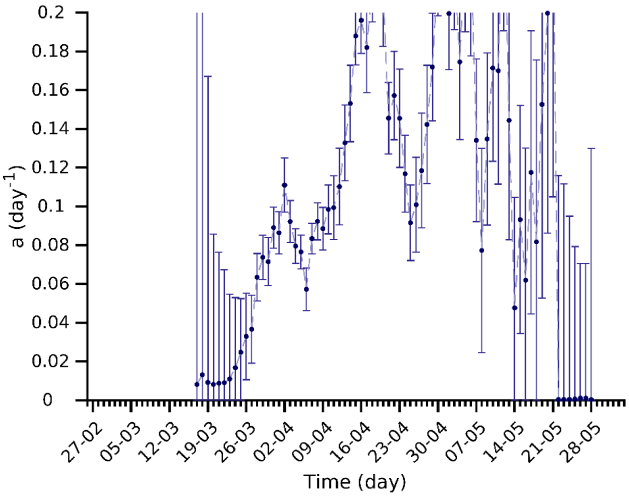


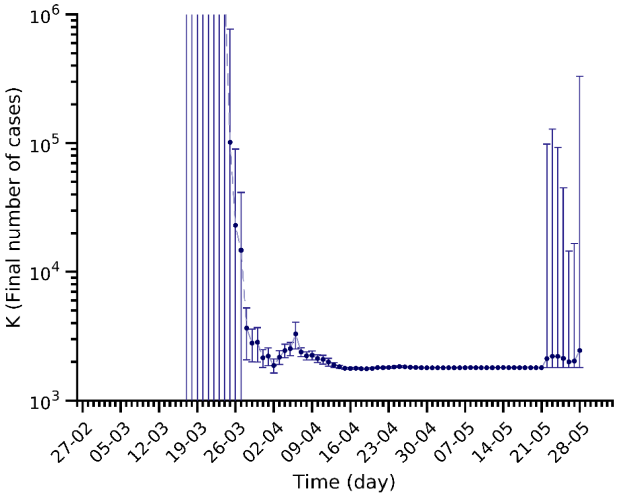


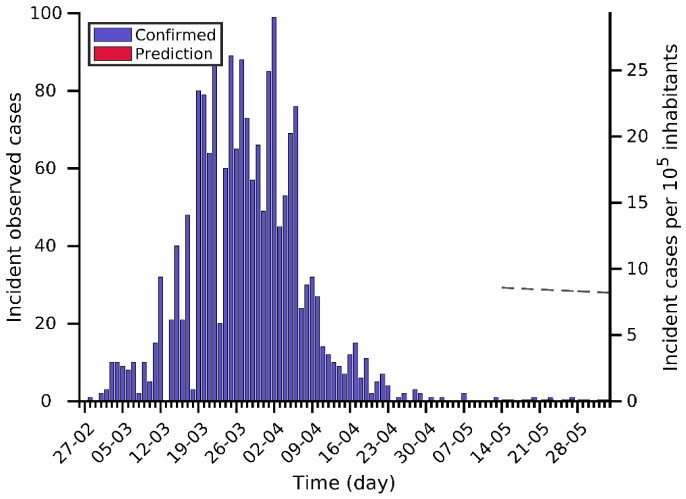
Iceland 28-05-2020. Population: 0.3M. Current cumulated incidence: $529/10^5$

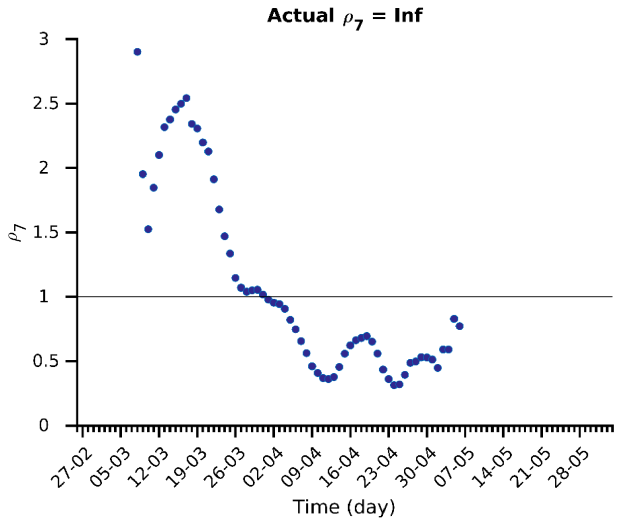


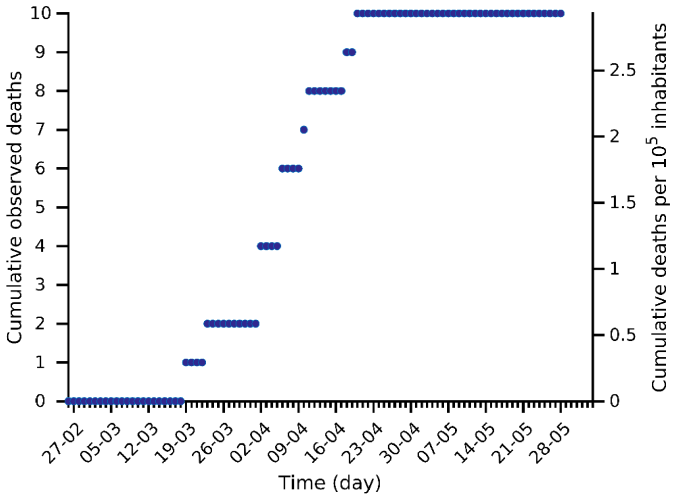


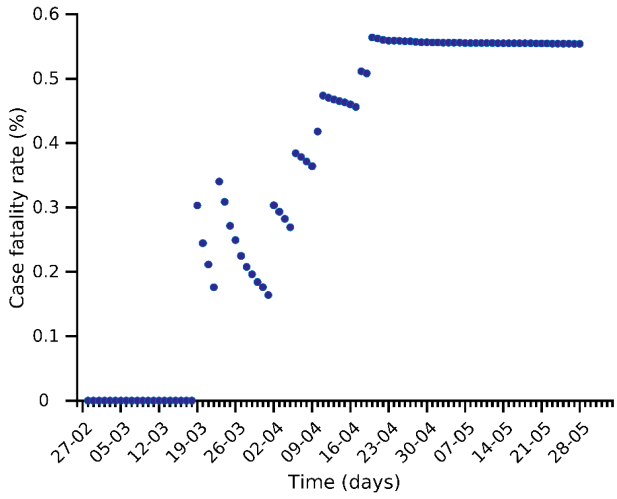




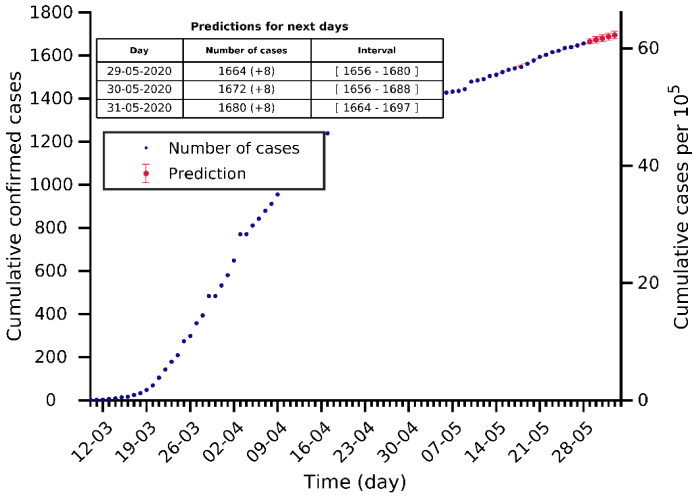


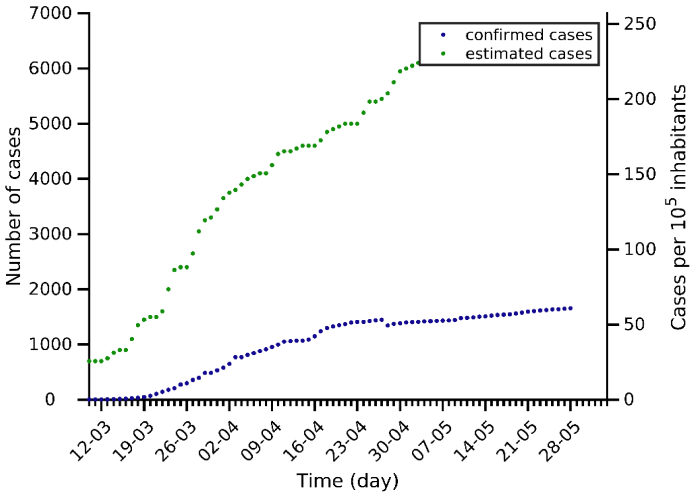


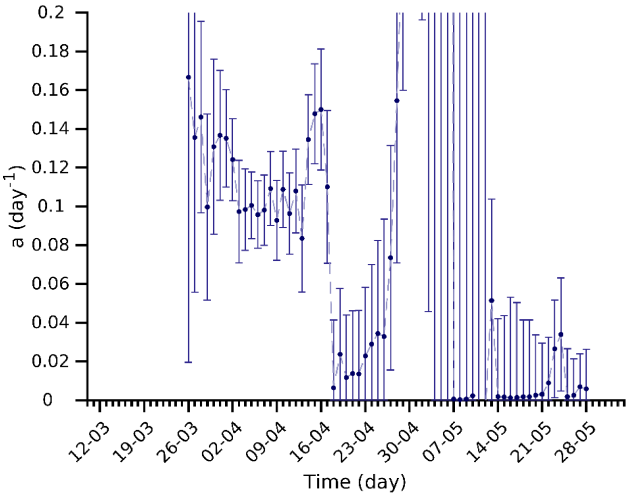


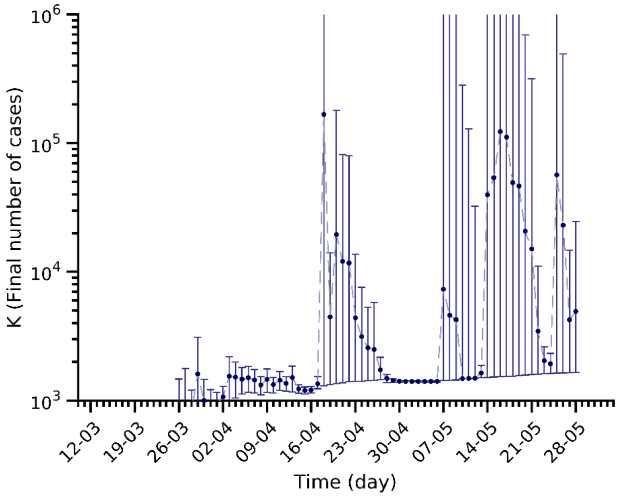


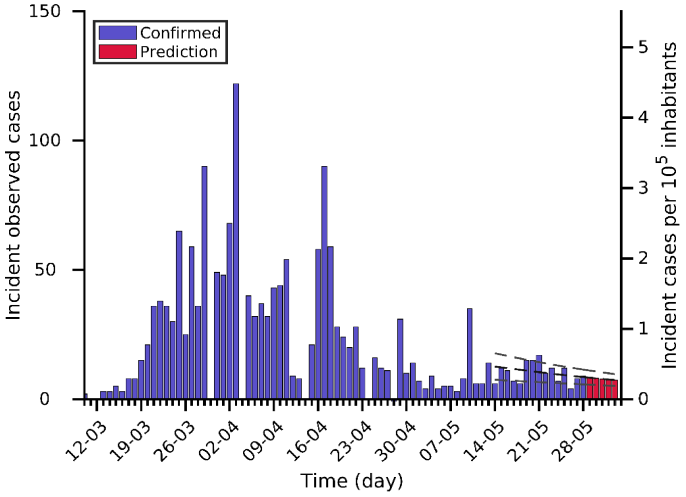
Lithuania 28-05-2020. Population: 2.7M. Current cumulated incidence: 61/10^5

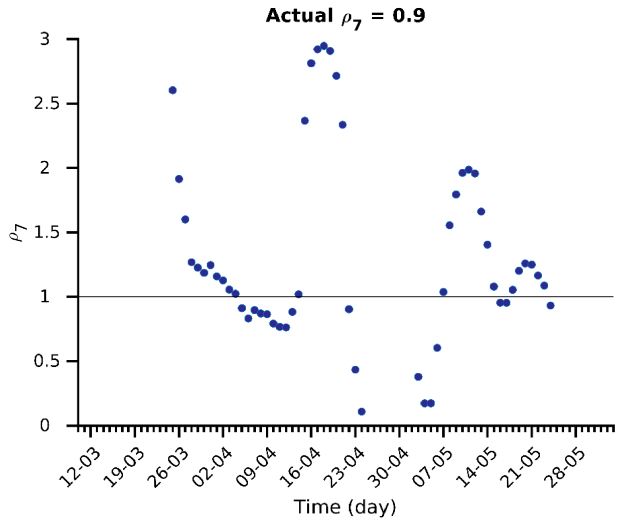


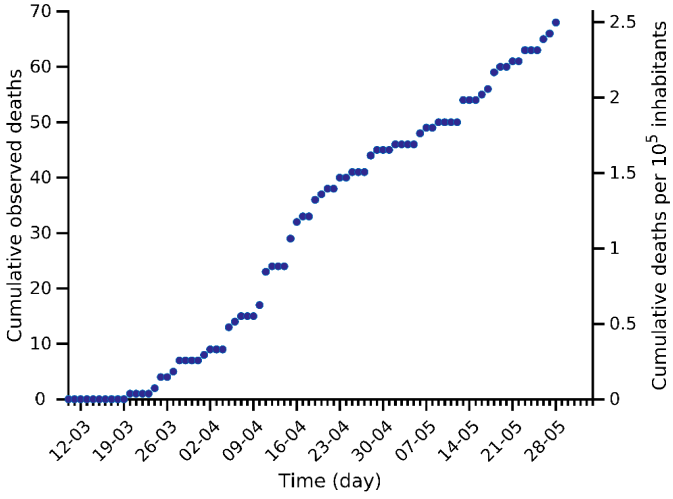


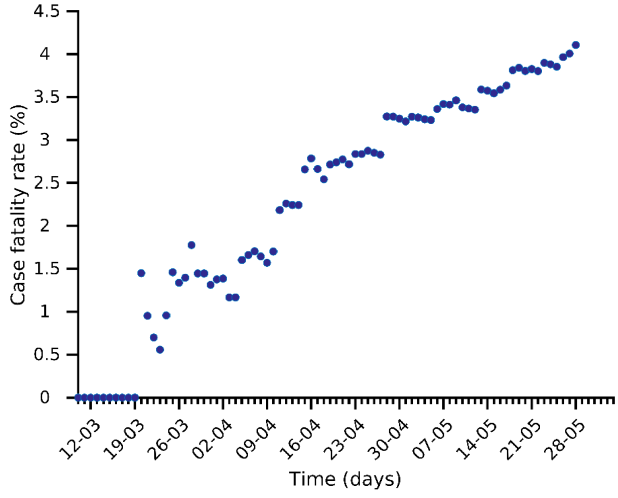




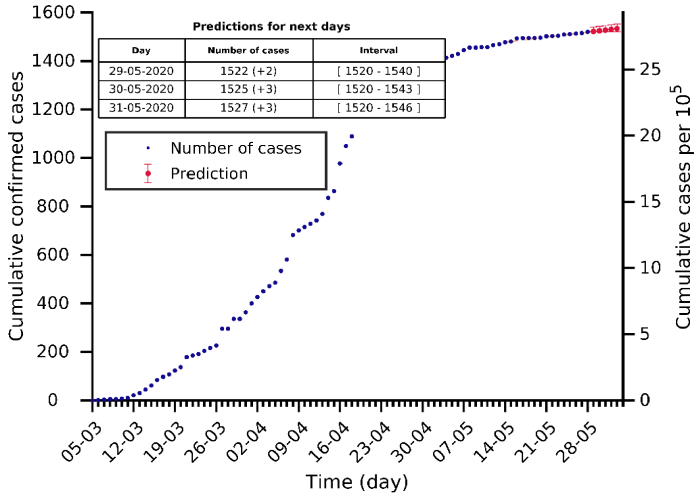


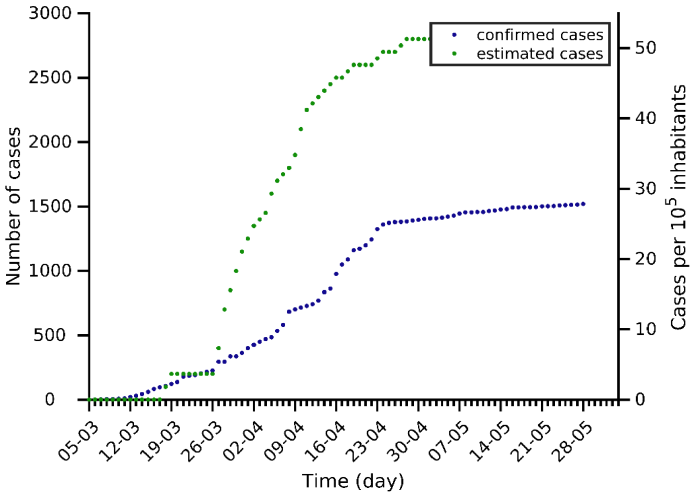


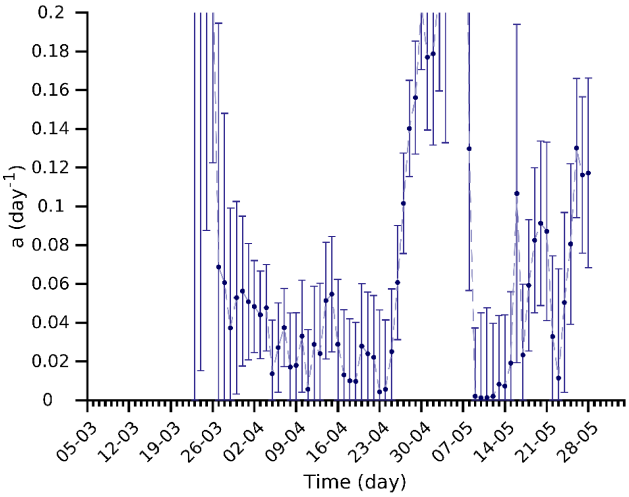


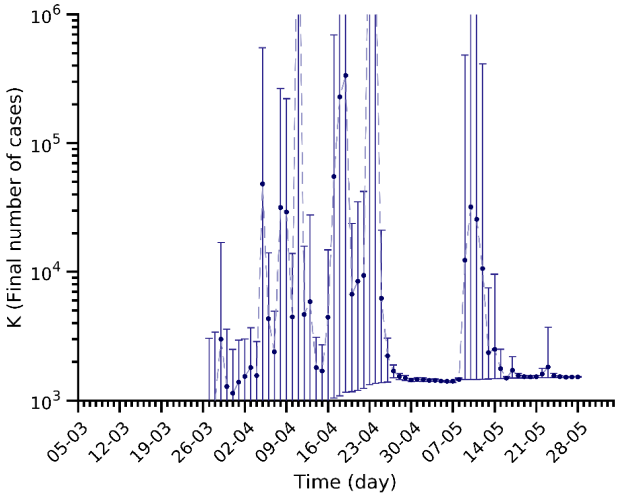


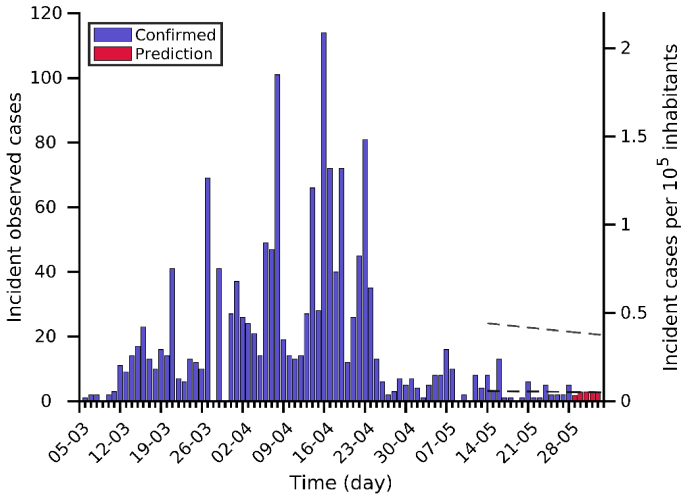
Slovakia 28-05-2020. Population: 5.5M. Current cumulated incidence: 28/10^5

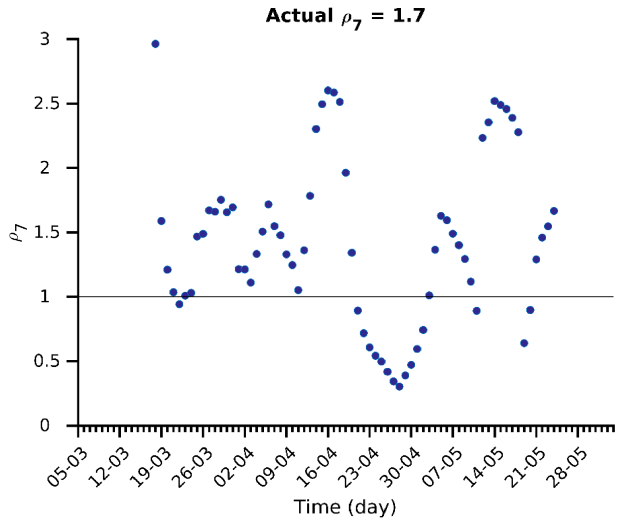


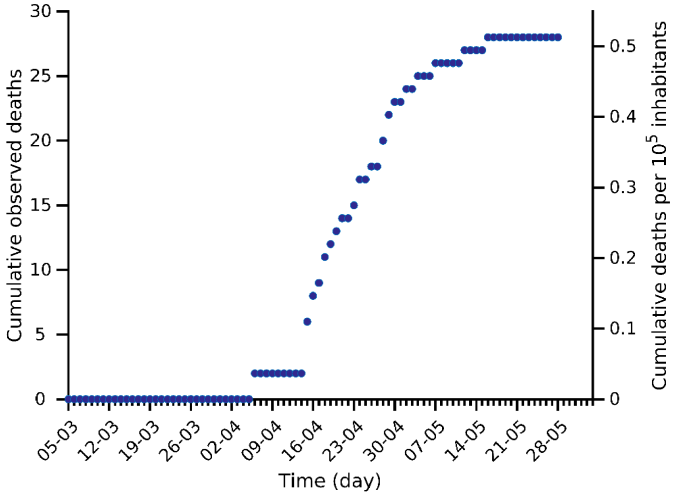


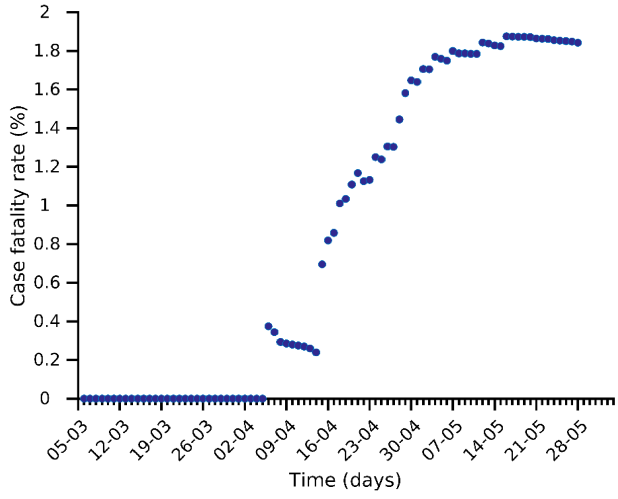




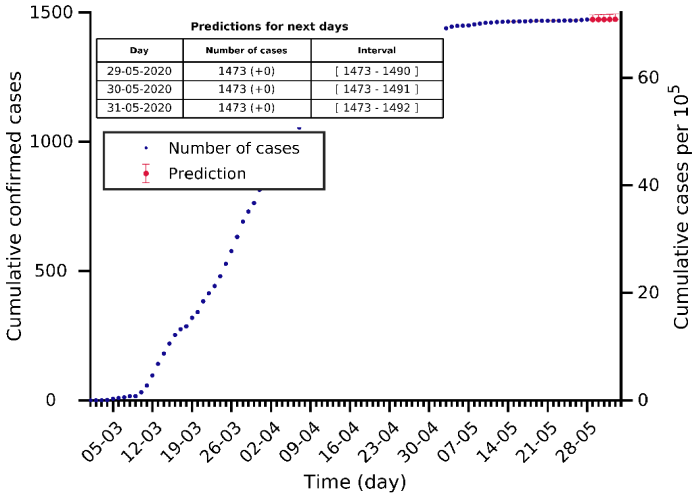


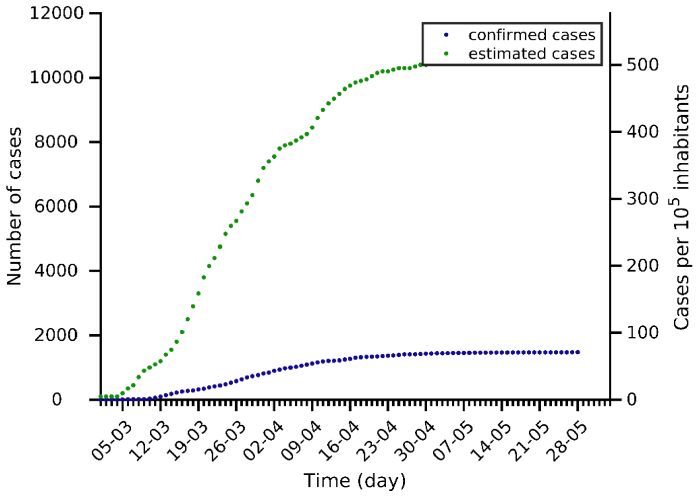


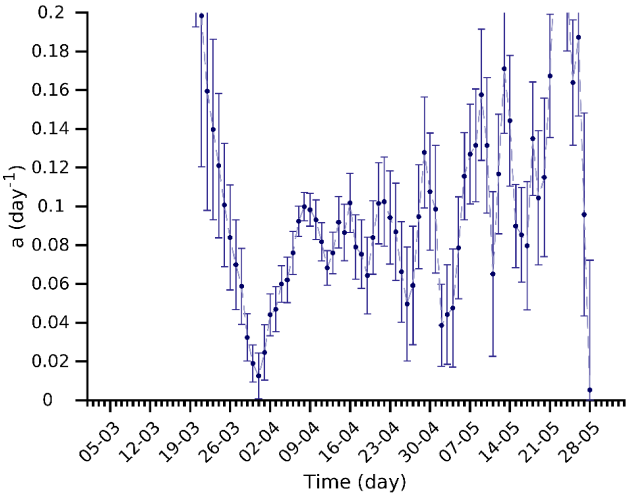


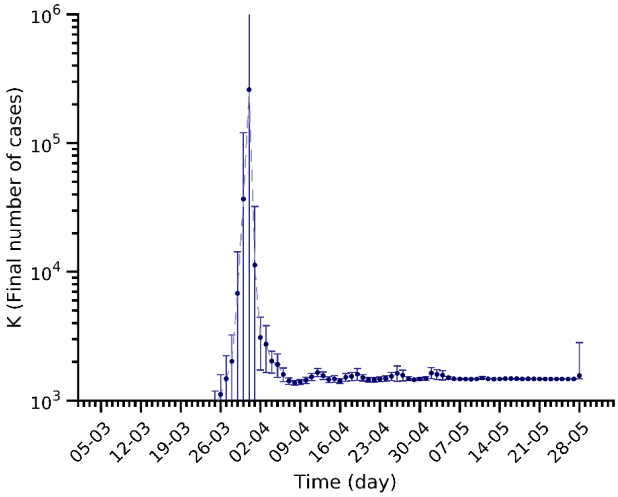


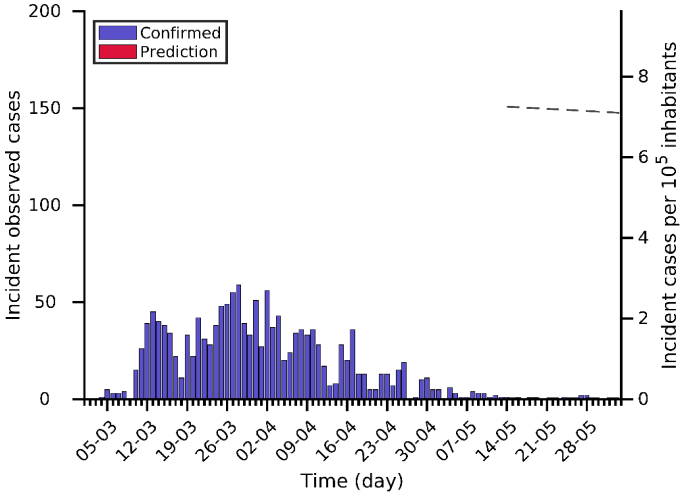
Slovenia 28-05-2020. Population: 2.1M. Current cumulated incidence: 71/10^5

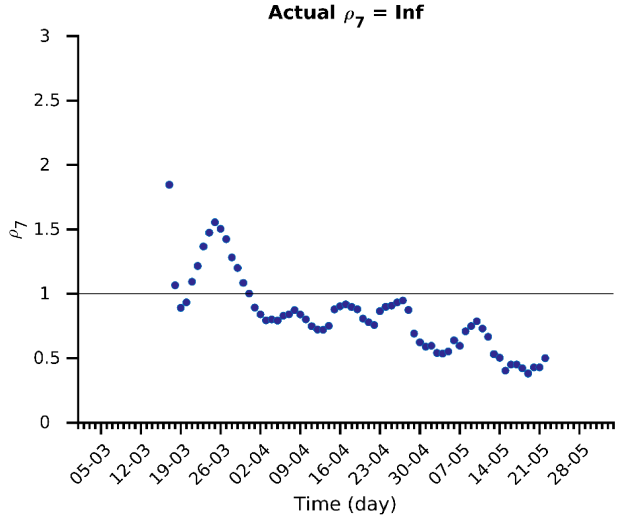


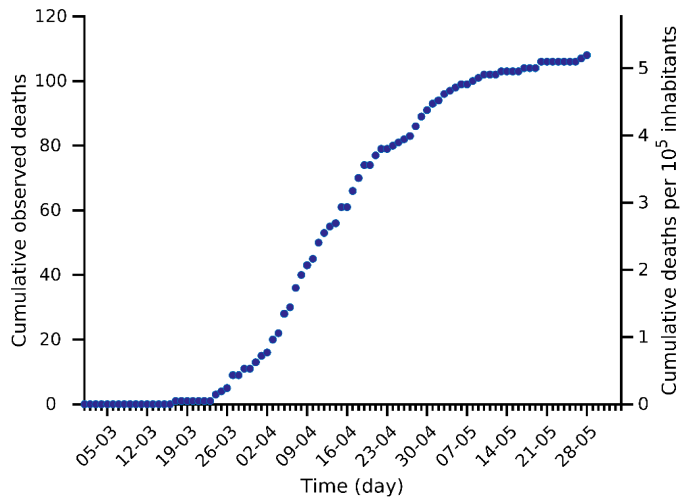


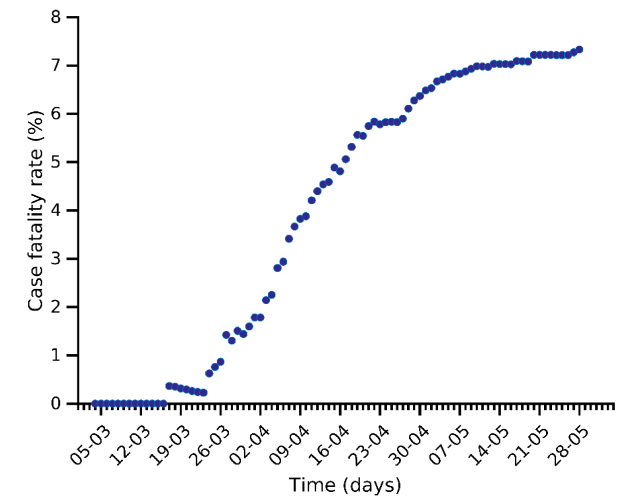




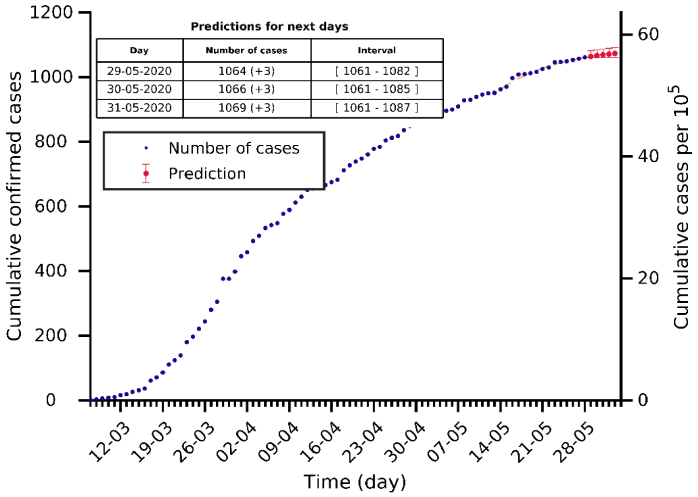


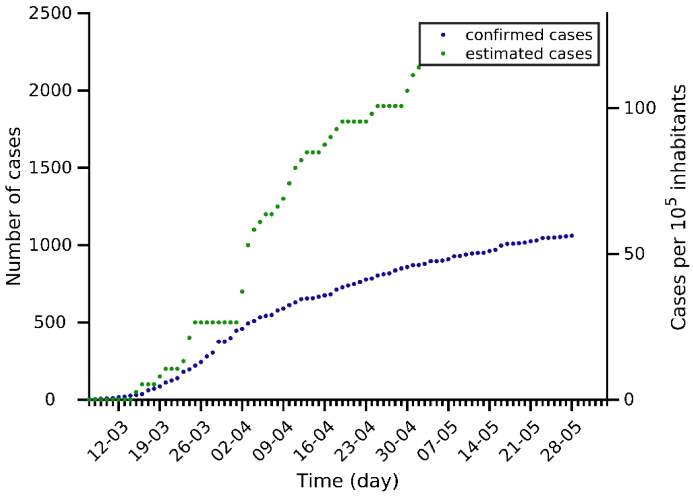


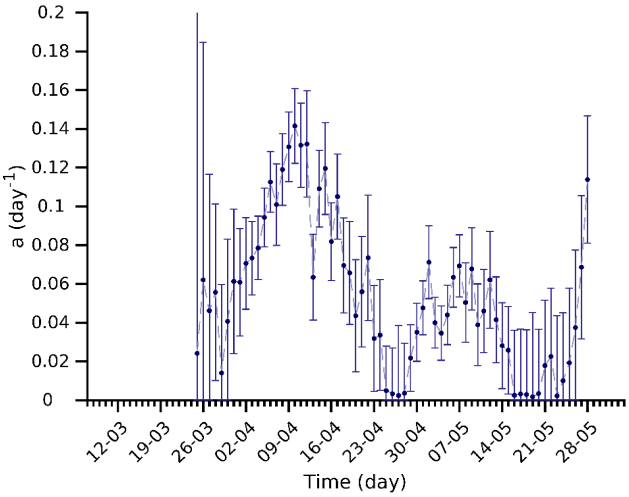


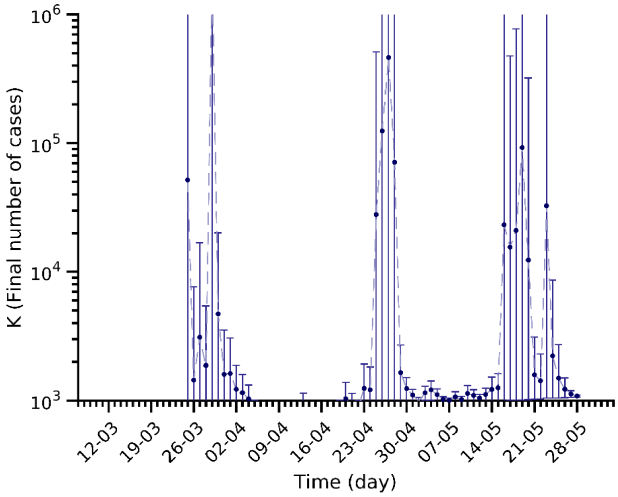


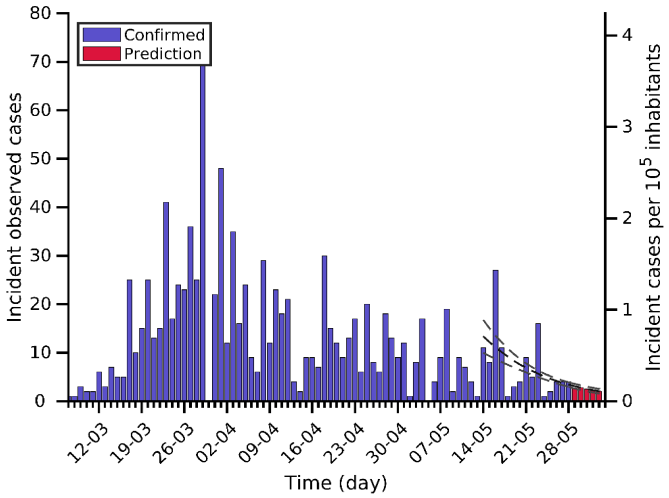
Latvia 28-05-2020. Population: 1.9M. Current cumulated incidence: $56/10^5$

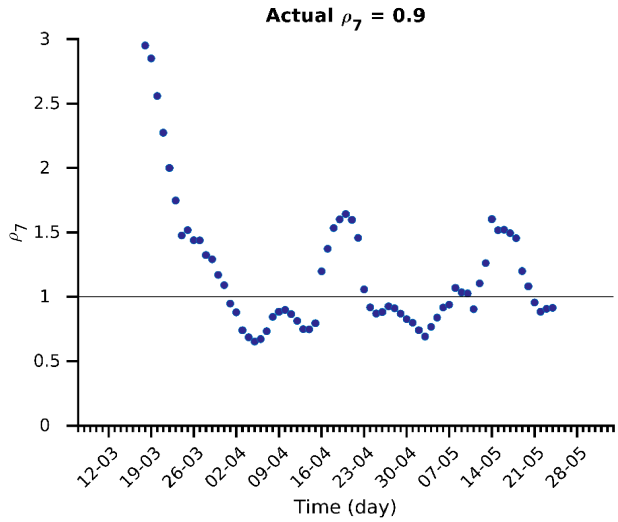


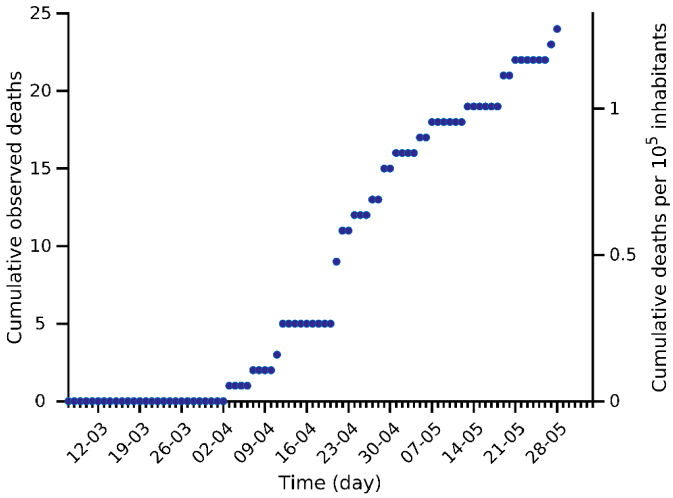


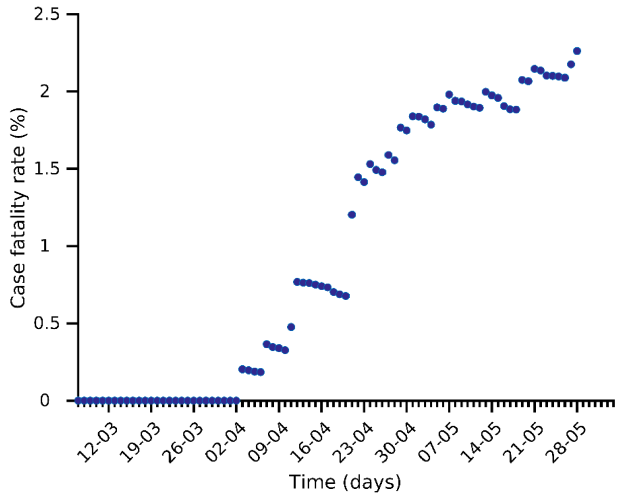




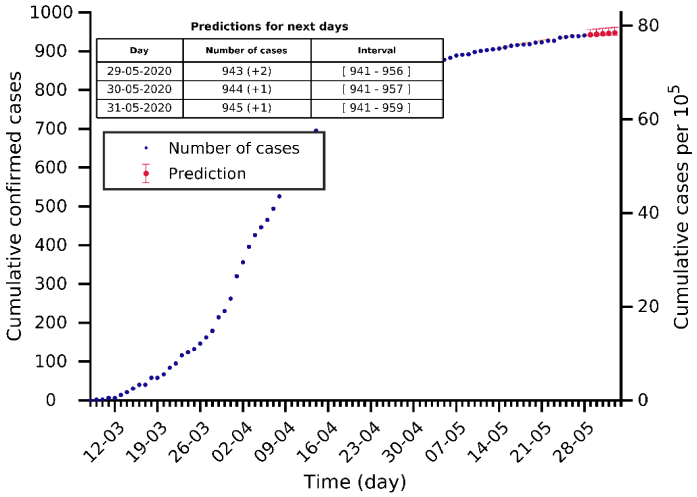


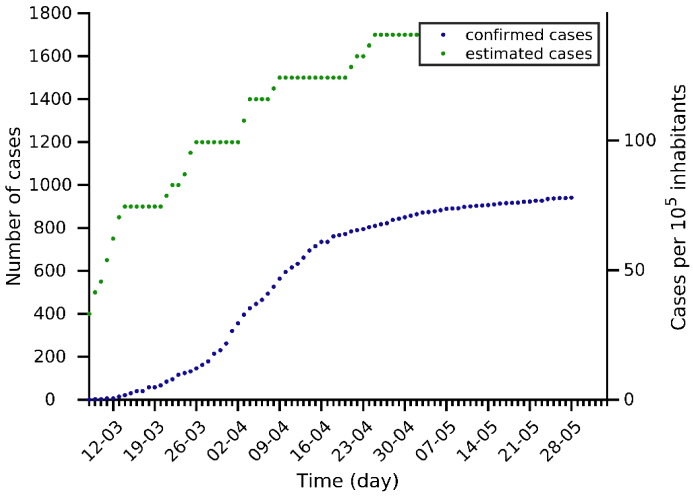


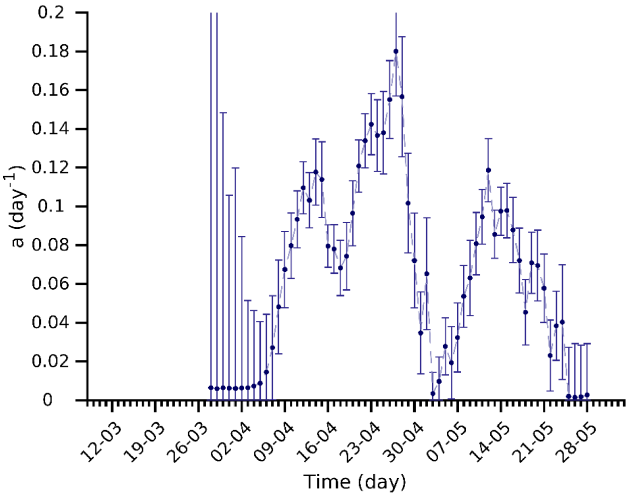


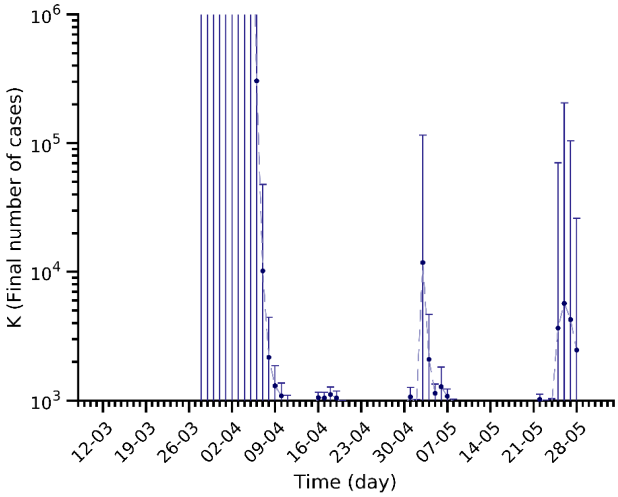


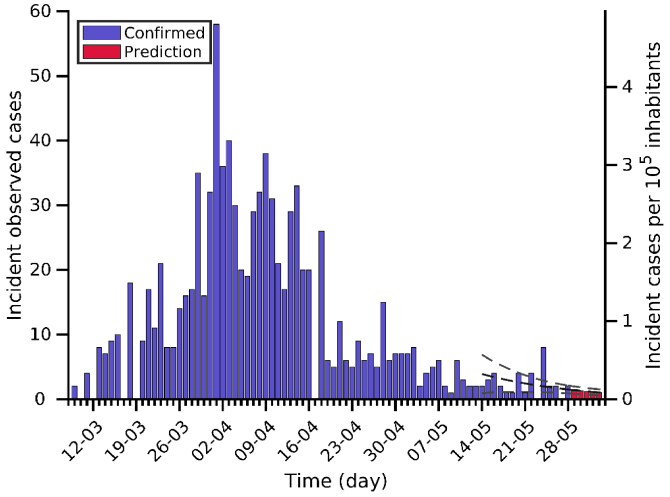
Cyprus 28-05-2020. Population: 1.2M. Current cumulated incidence: $78/10^5$

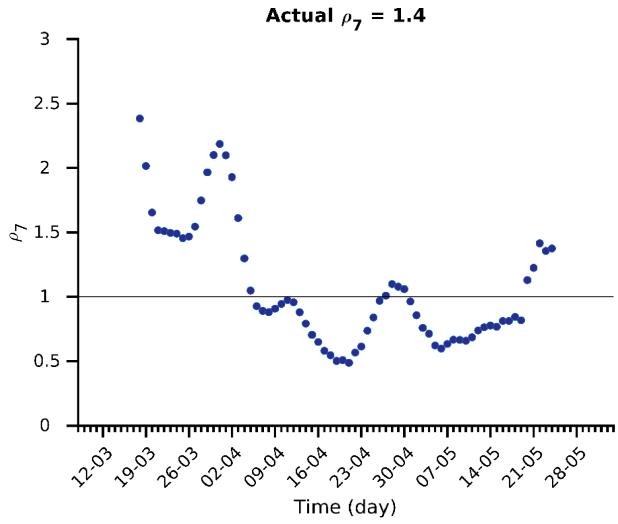


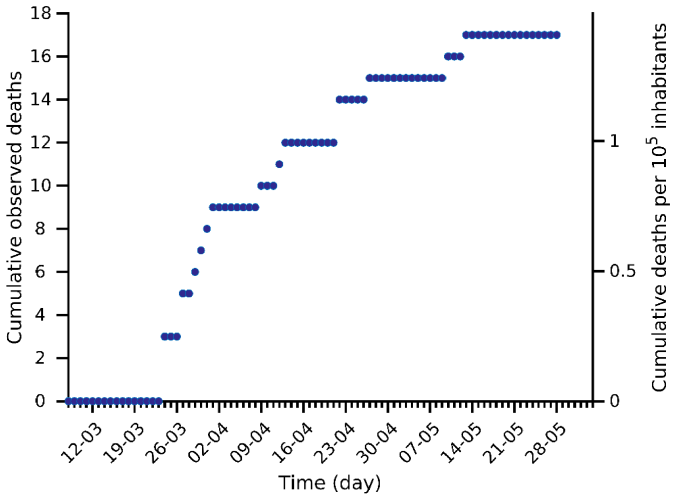


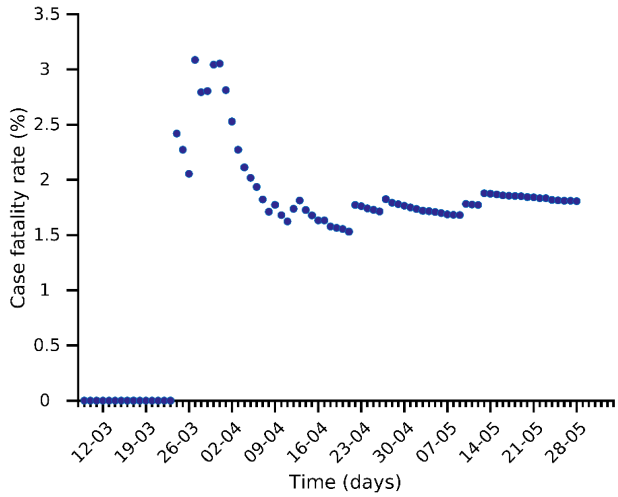




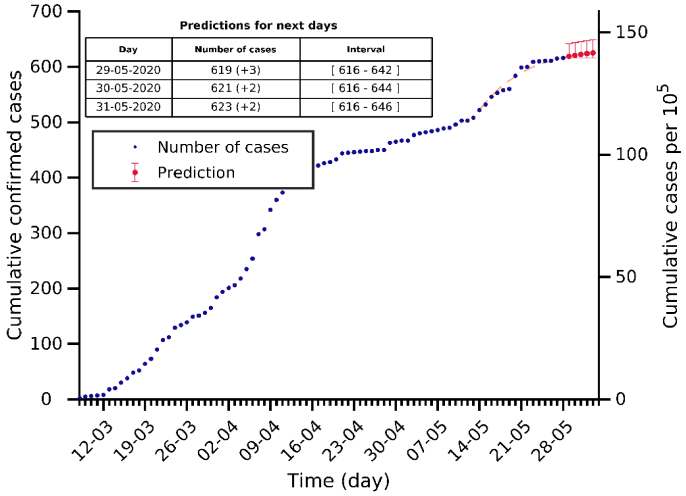


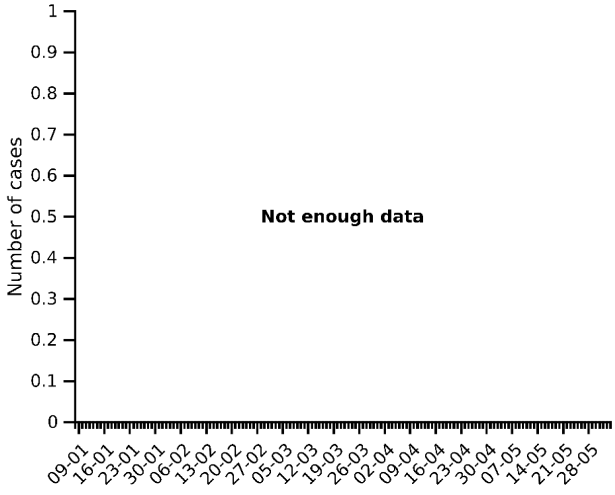


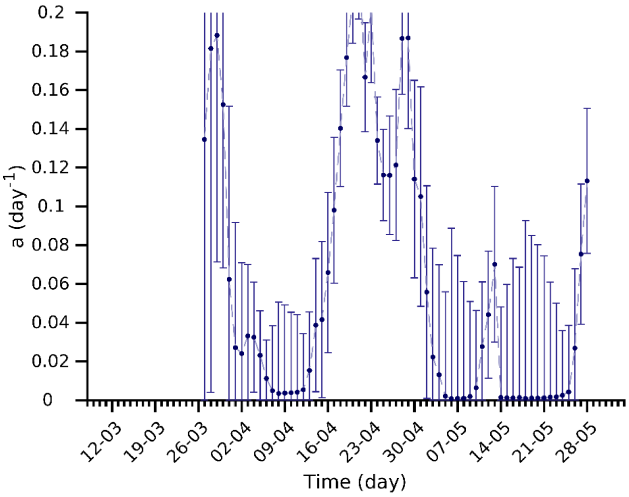


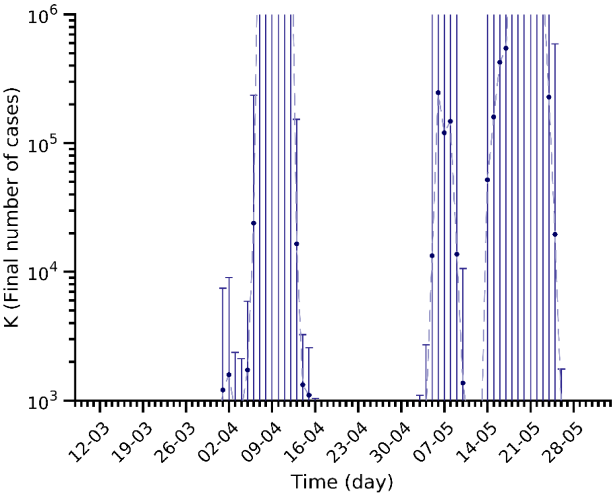


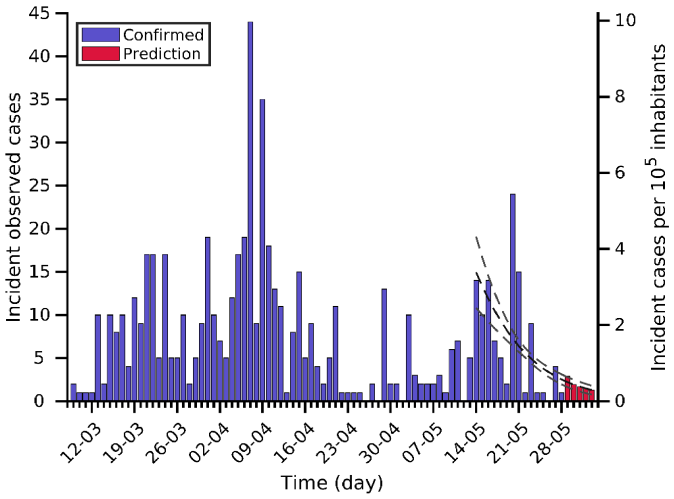
Malta 28-05-2020. Population: 0.4M. Current cumulated incidence: $140/10^5$

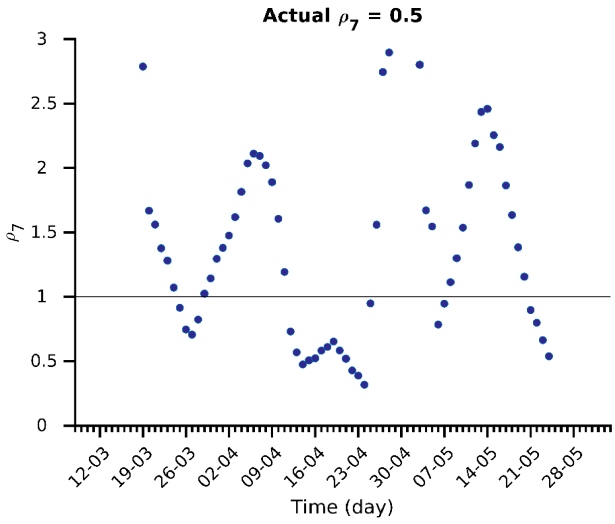


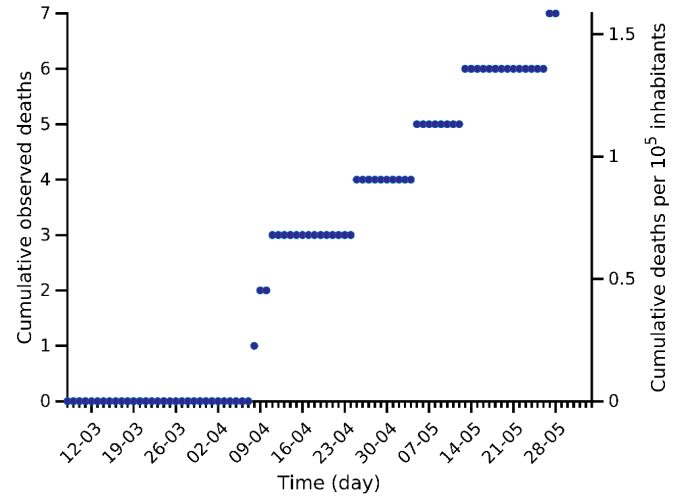


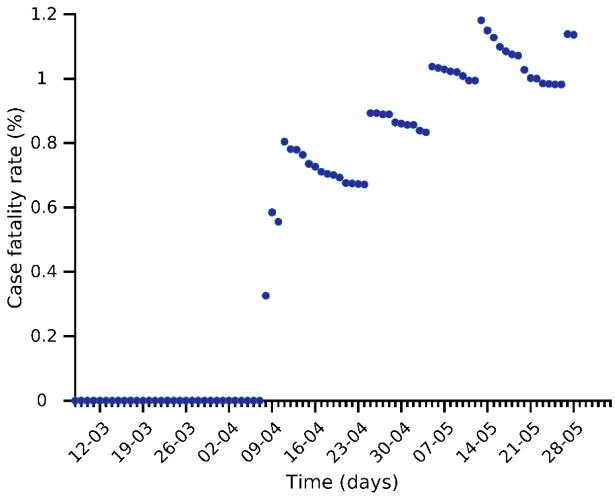








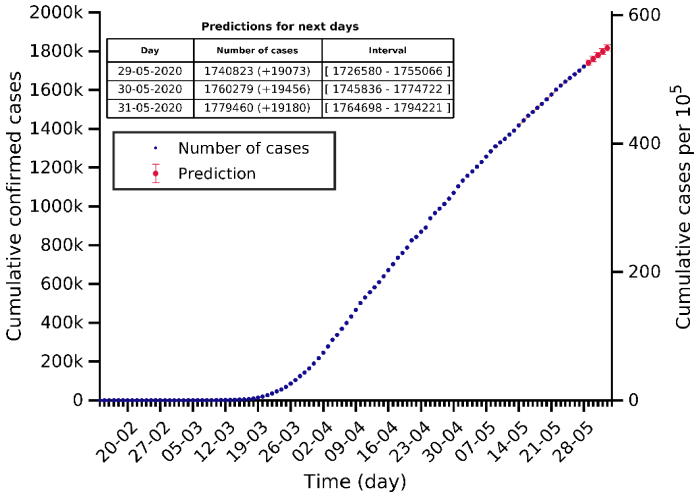


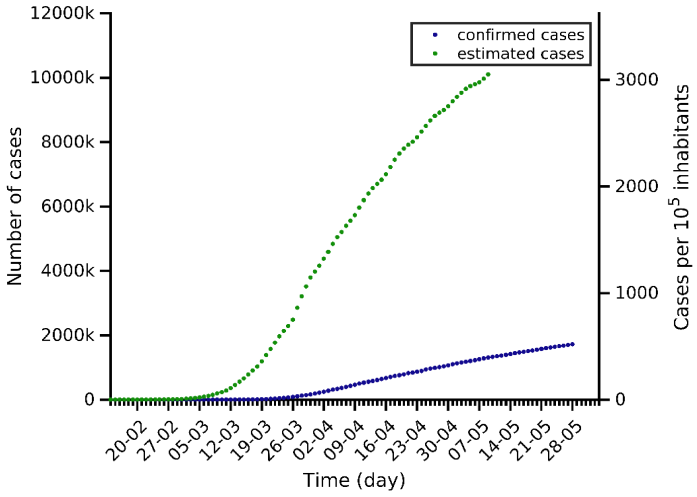


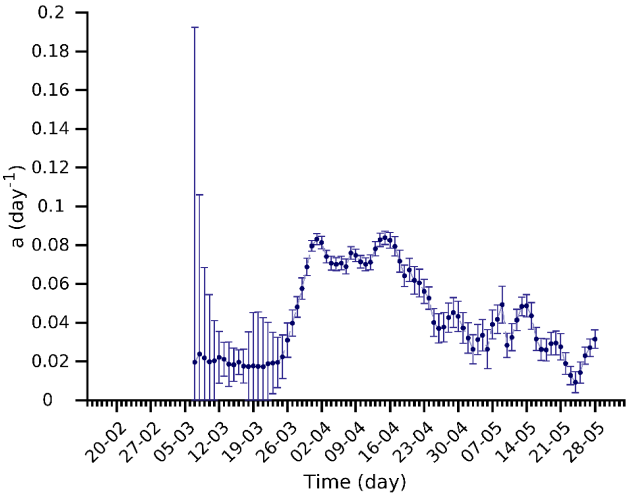
(2) Analysis and prediction of COVID-19 for other countries

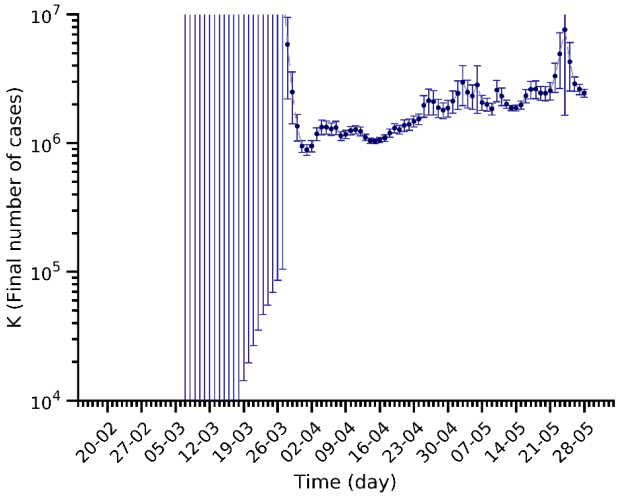
Data obtained from https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

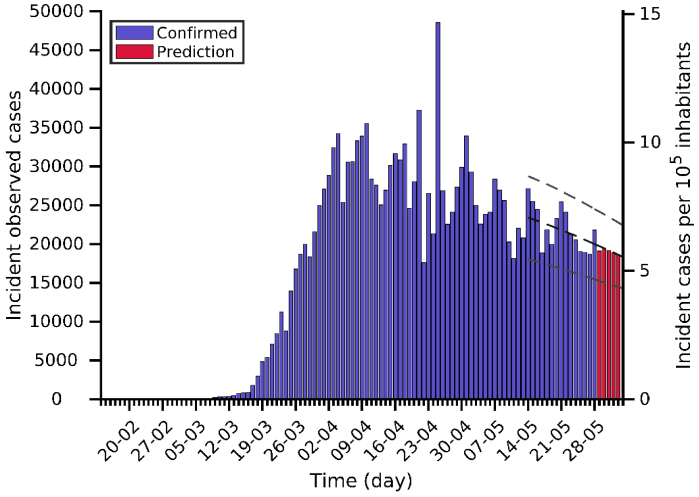
USA 28-05-2020. Population: 331.0M. Current cumulated incidence: 520/10^5

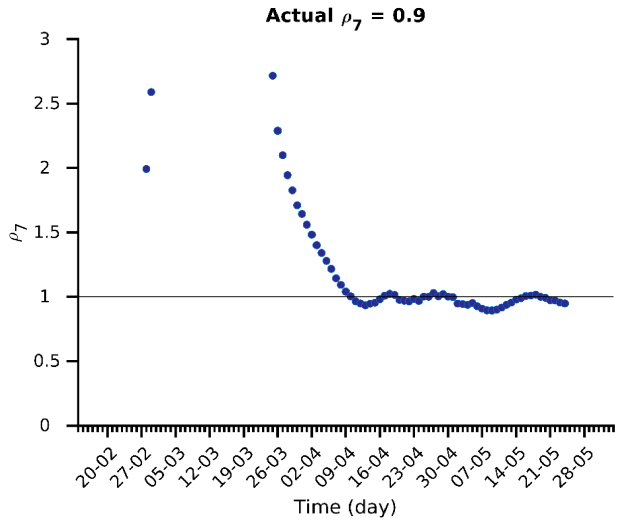


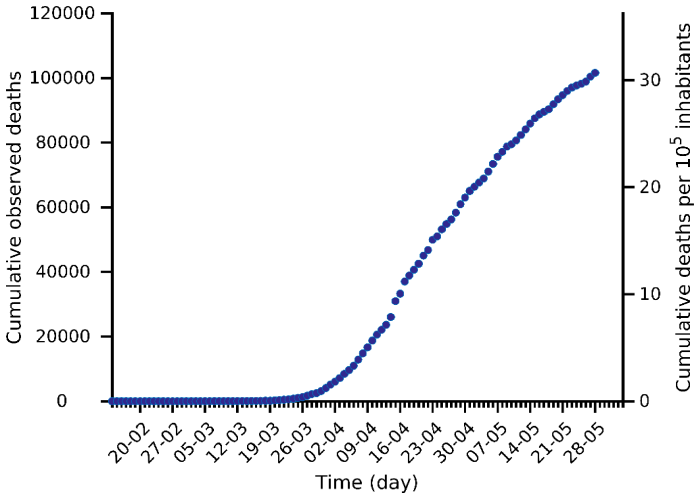


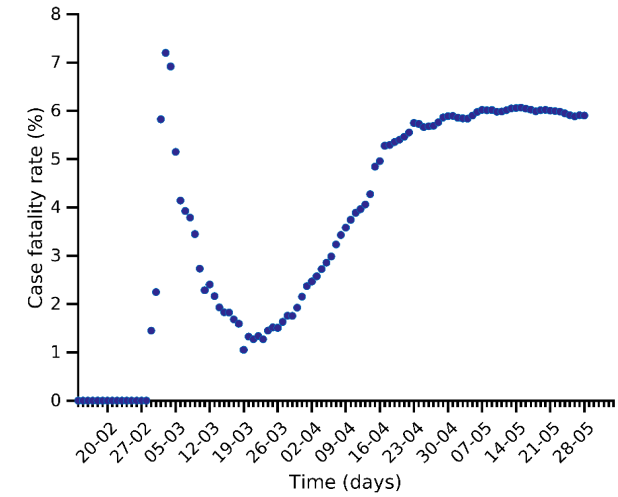




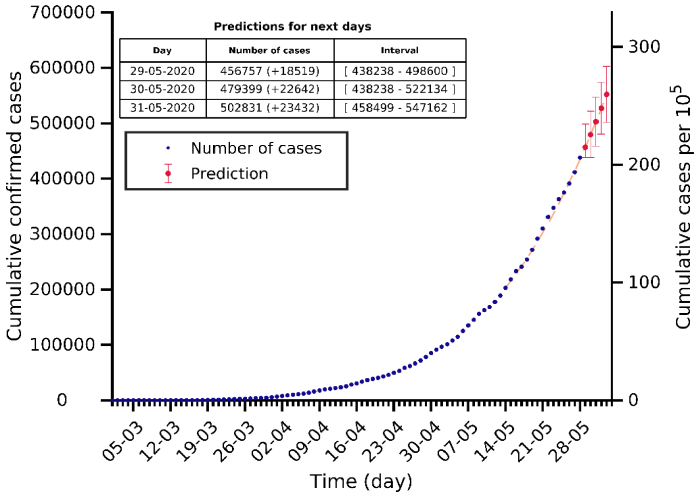


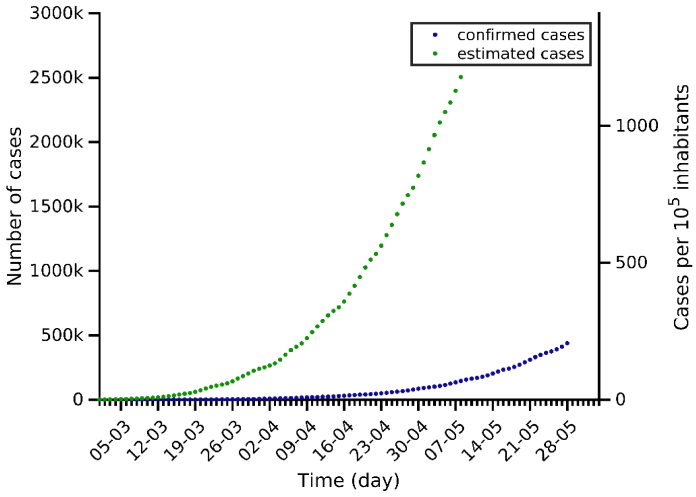


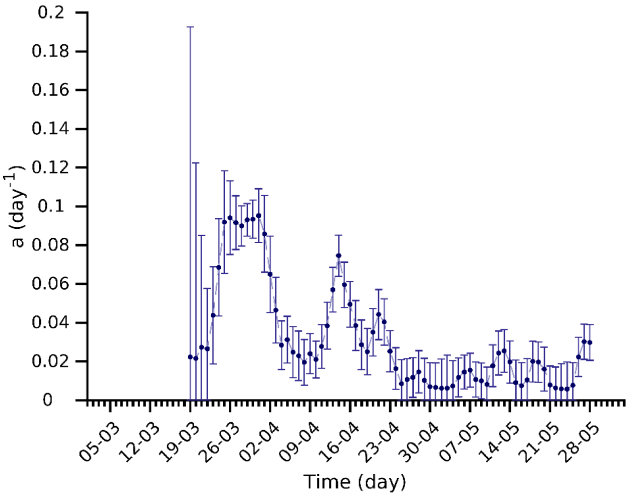


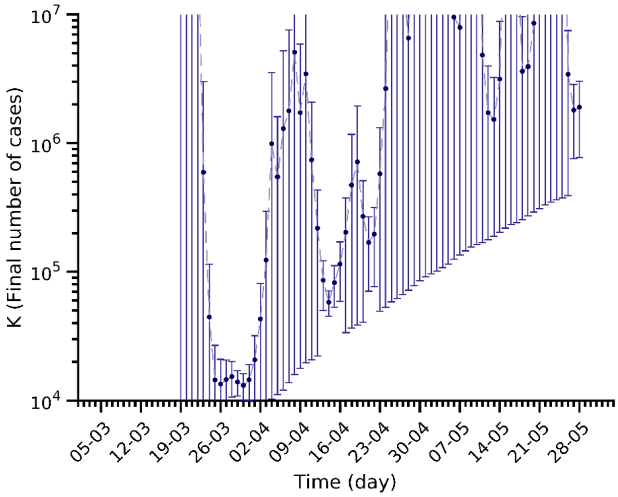


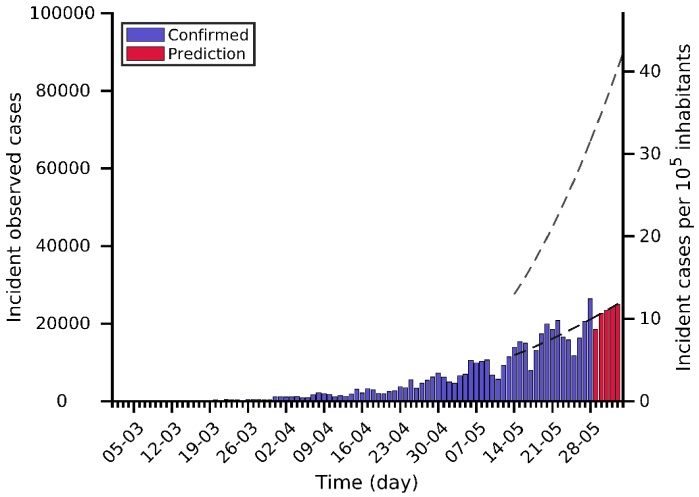
Brazil 28-05-2020. Population: 212.6M. Current cumulated incidence: 206/10^5

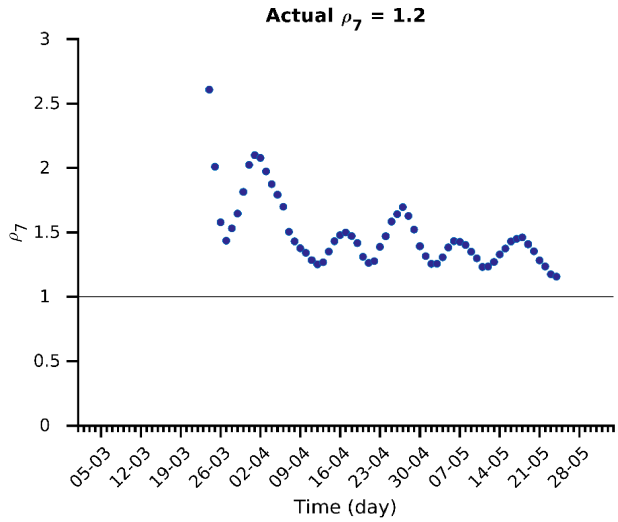


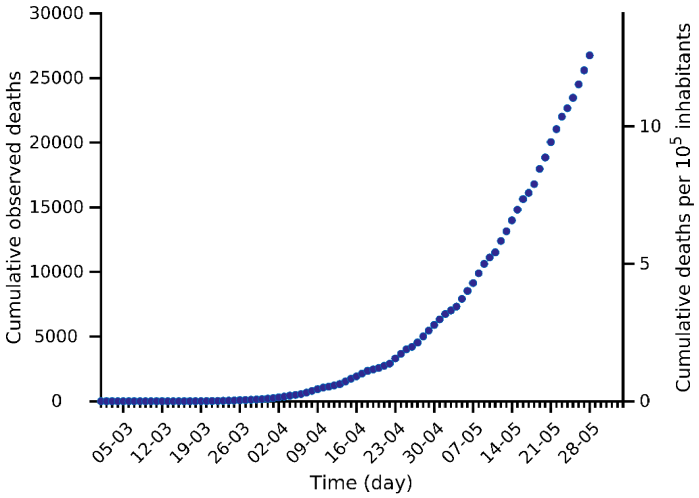


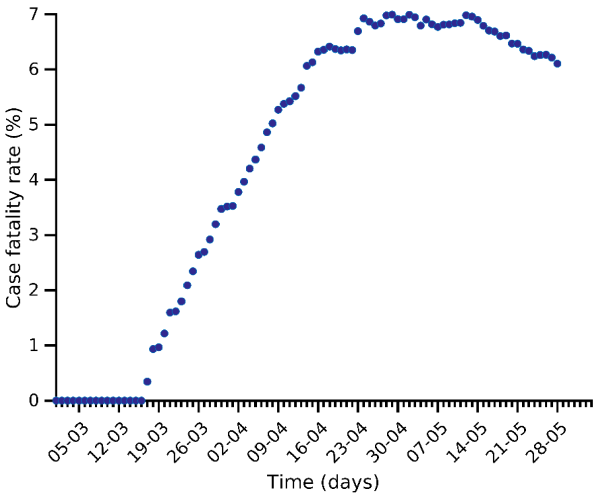




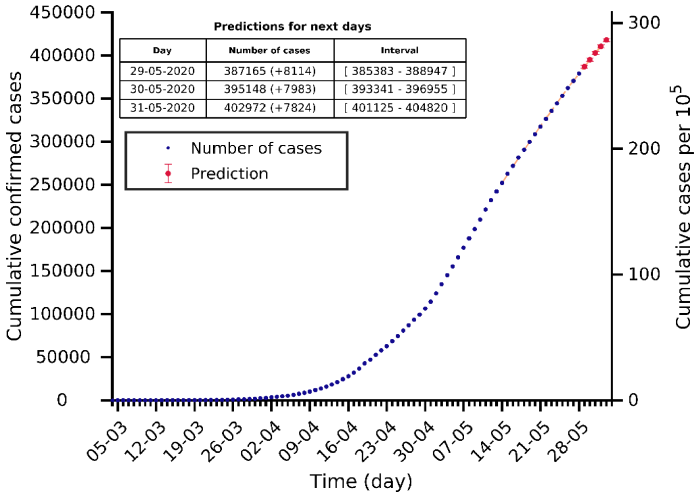


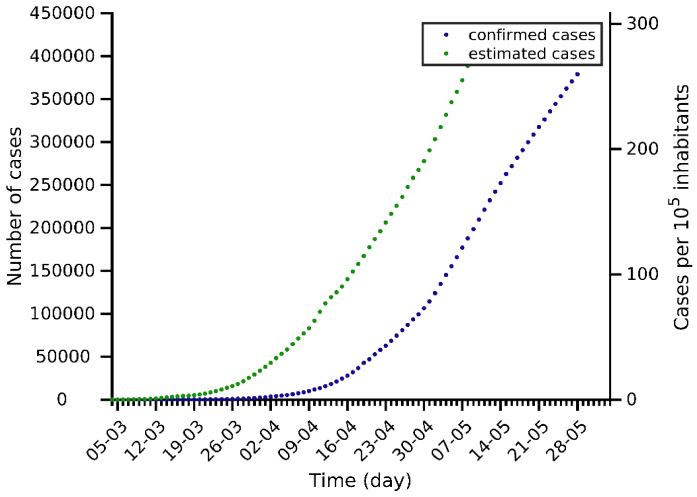


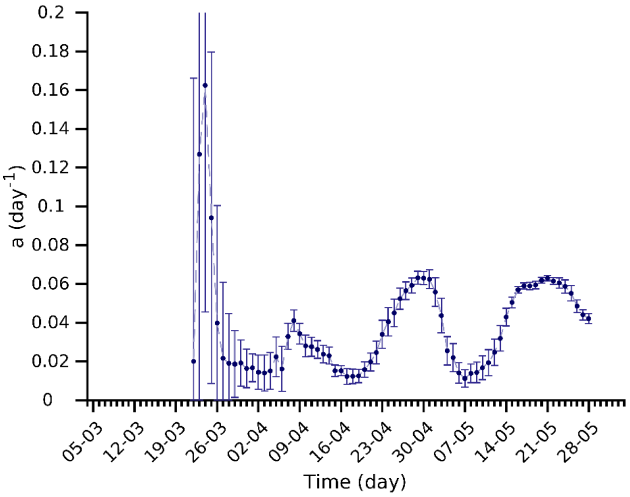


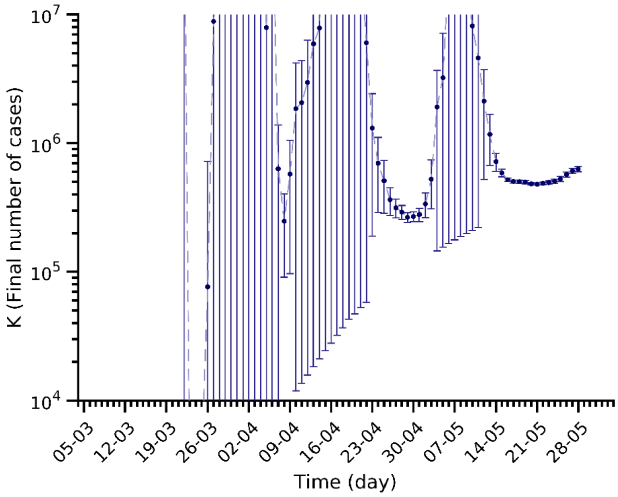


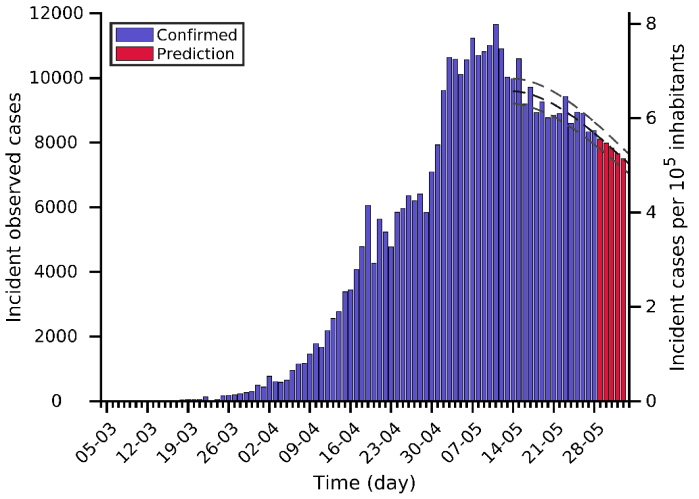
Russia 28-05-2020. Population: 145.9M. Current cumulated incidence: $260/10^5$

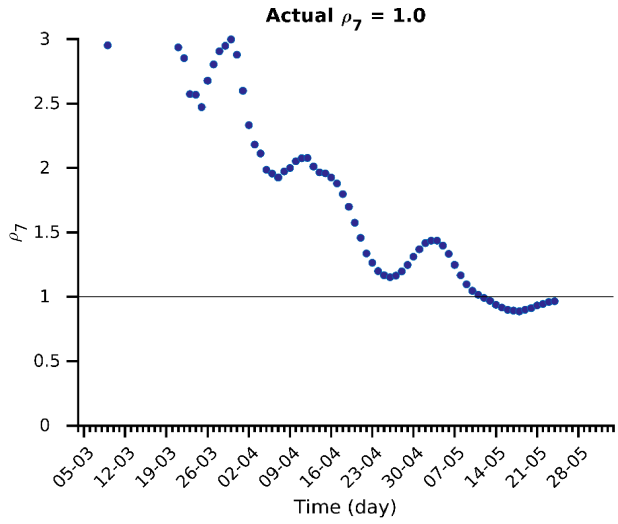


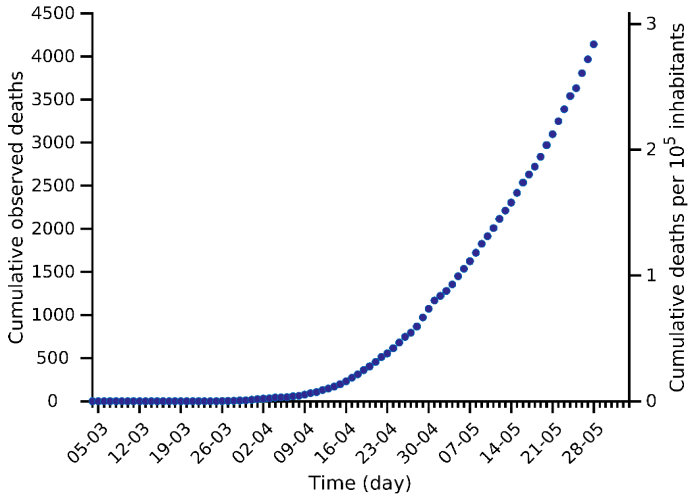


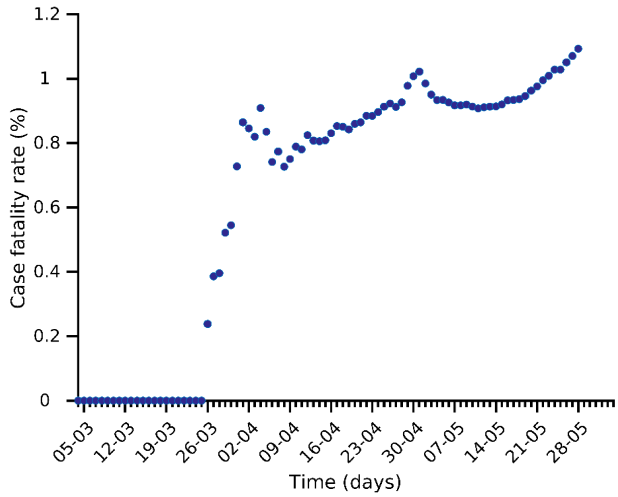




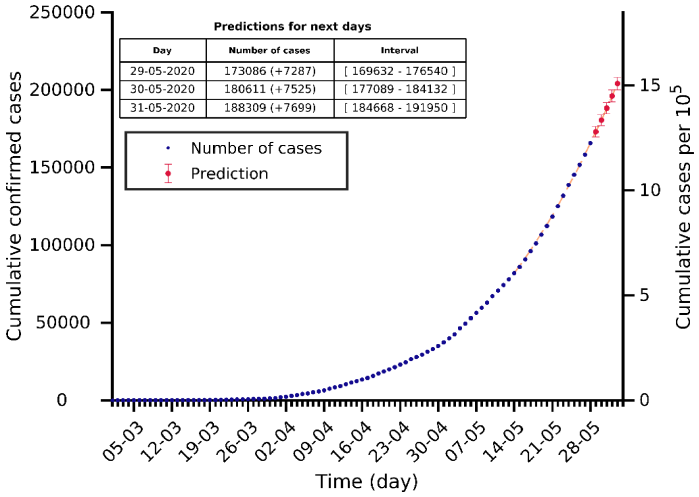


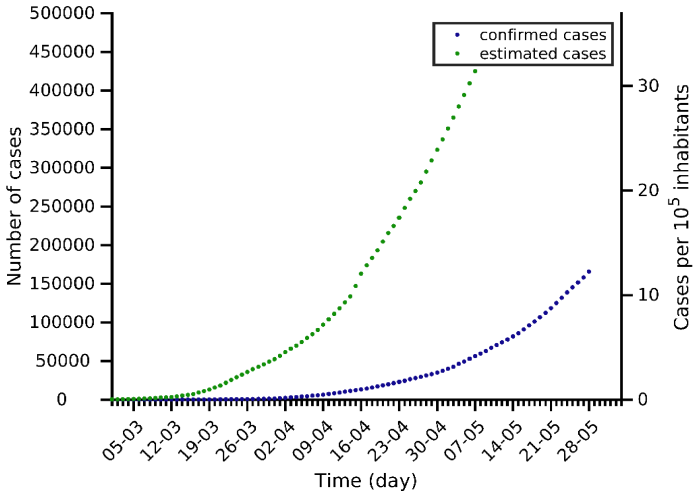


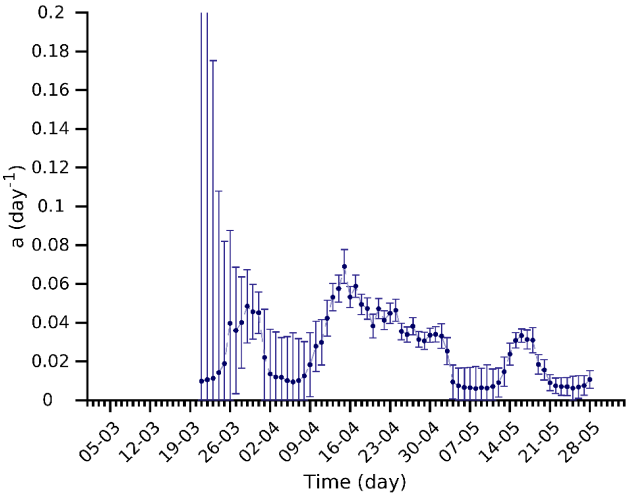


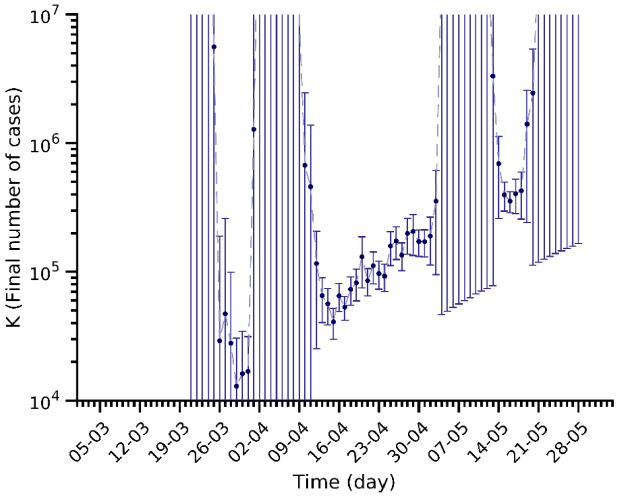


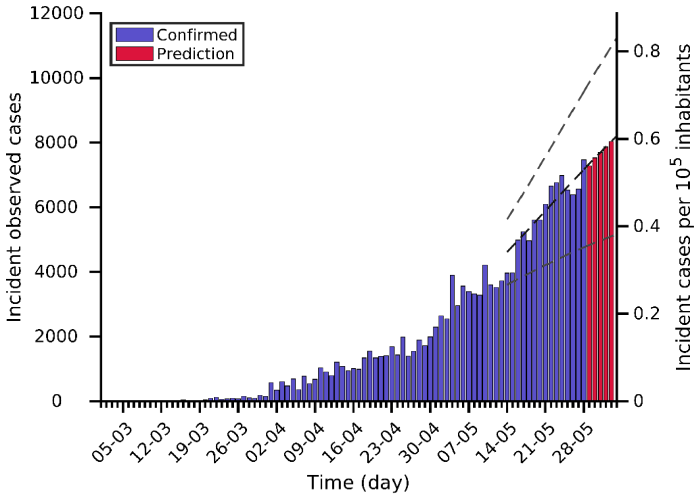
India 28-05-2020. Population: 1353.0M. Current cumulated incidence: $12/10^5$

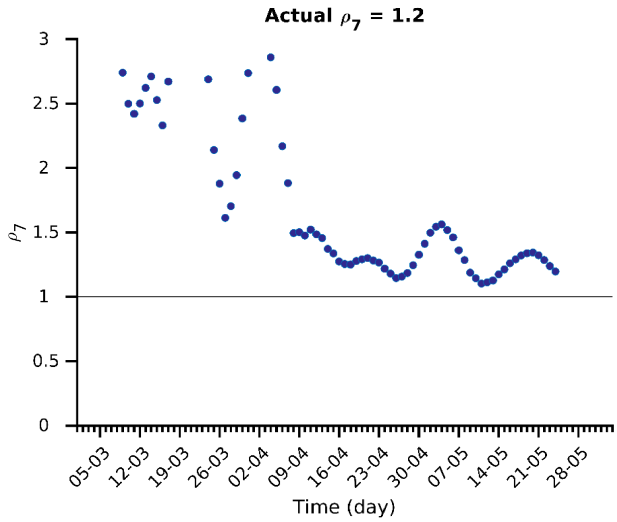


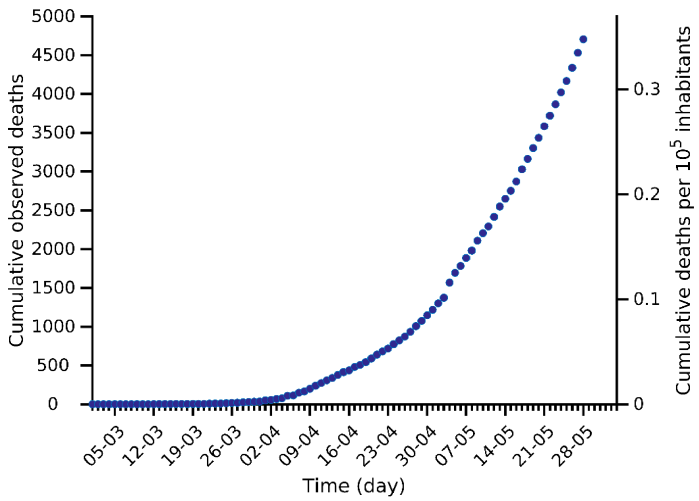


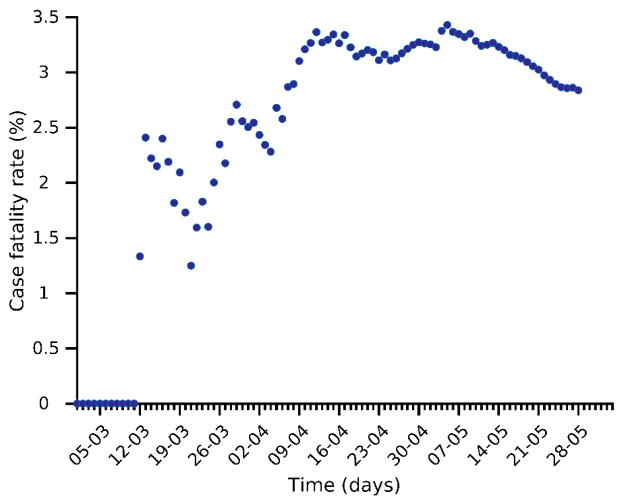




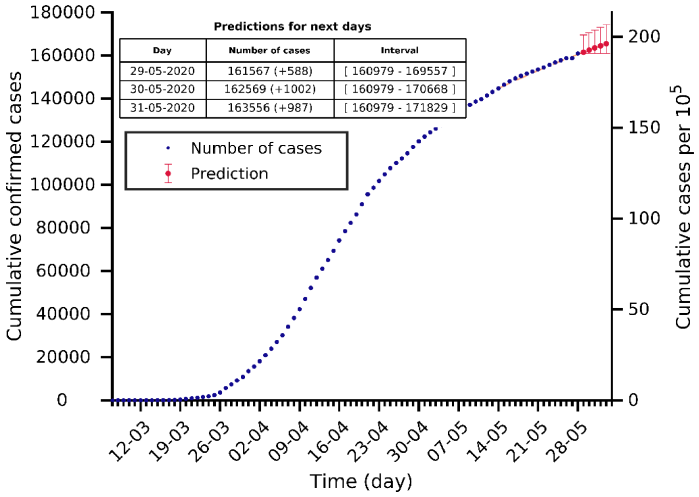


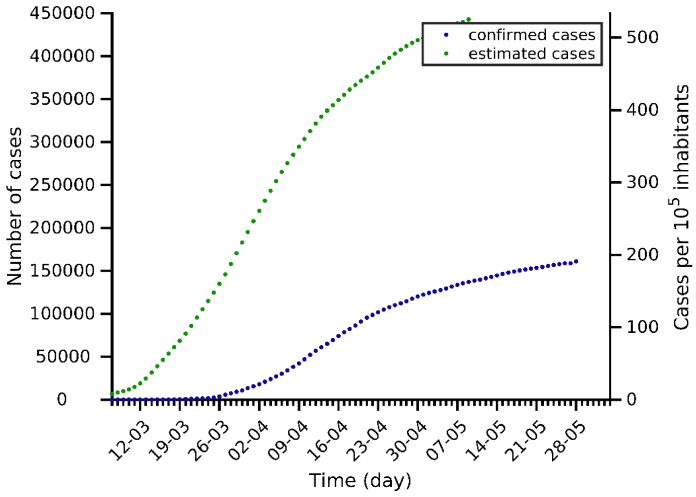


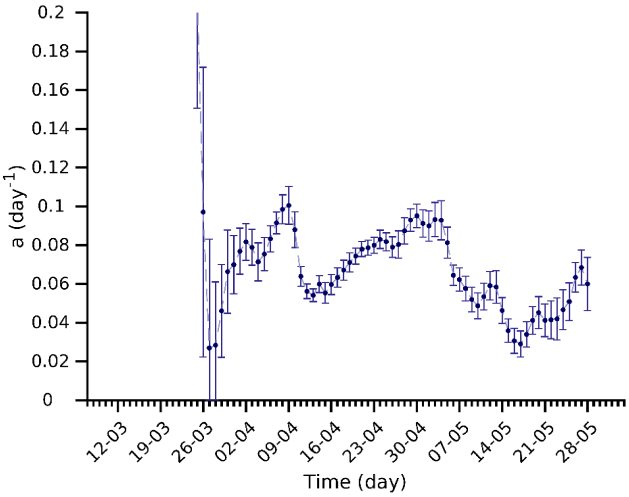


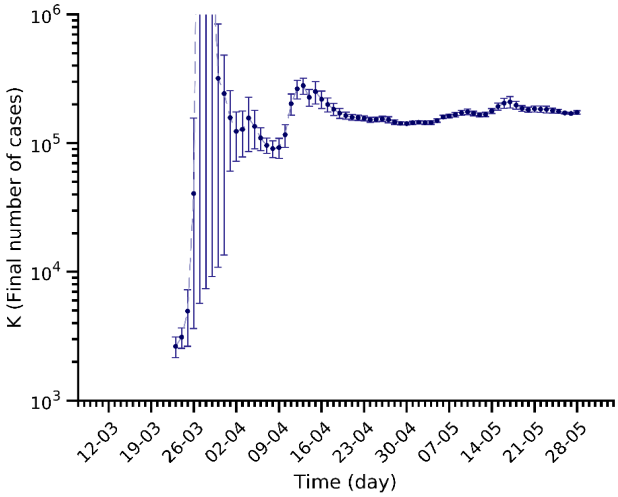


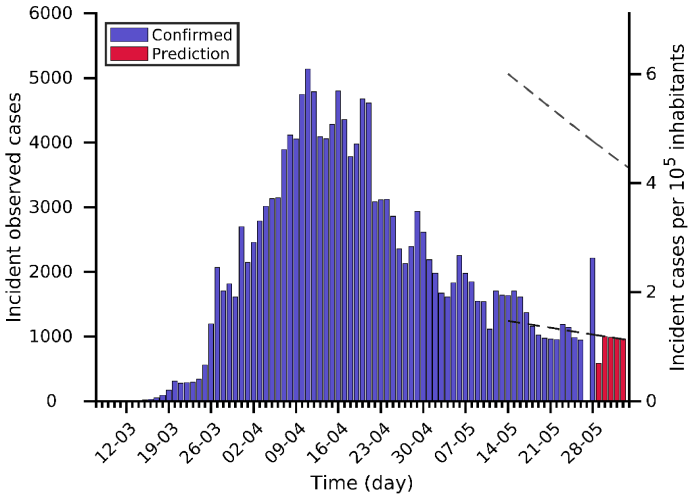
Turkey 28-05-2020. Population: 84.3M. Current cumulated incidence: 191/10^5

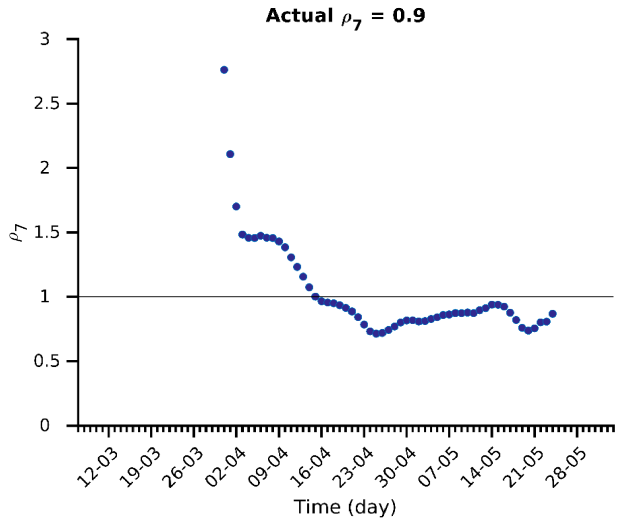


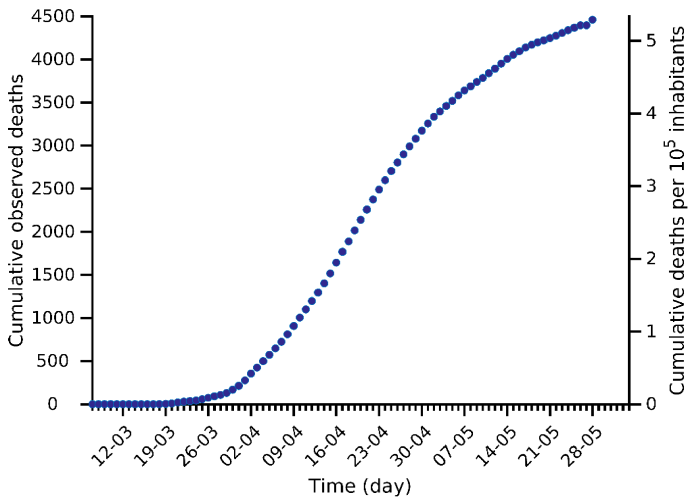


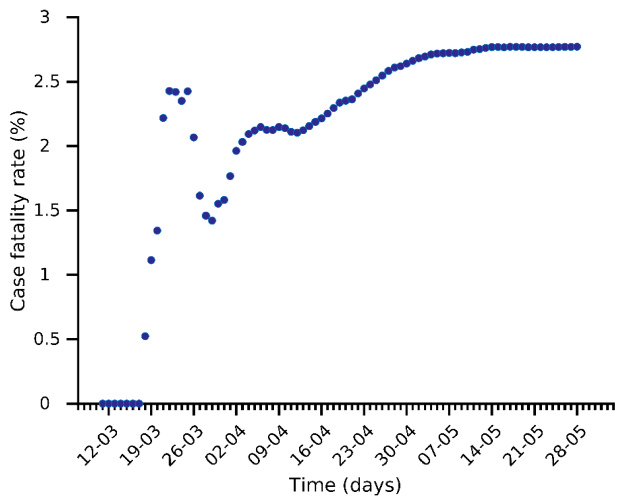




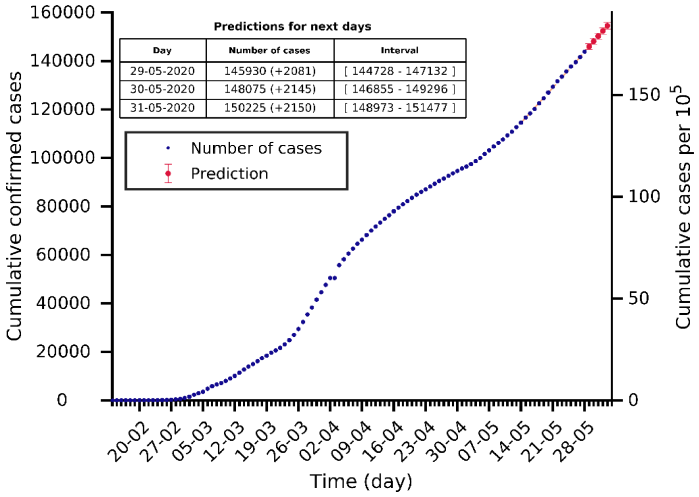


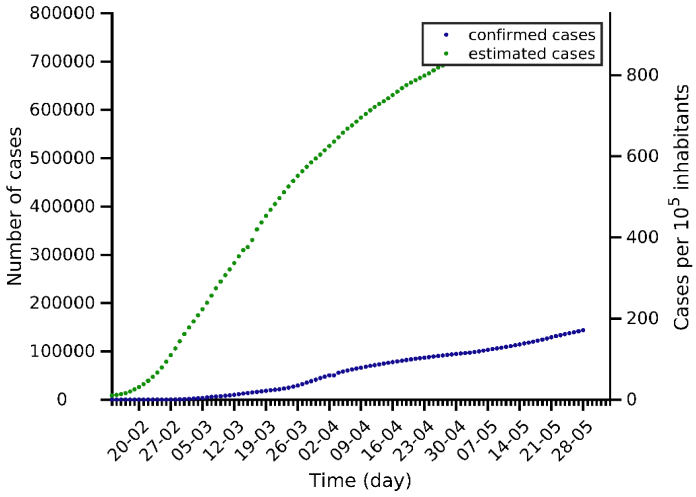


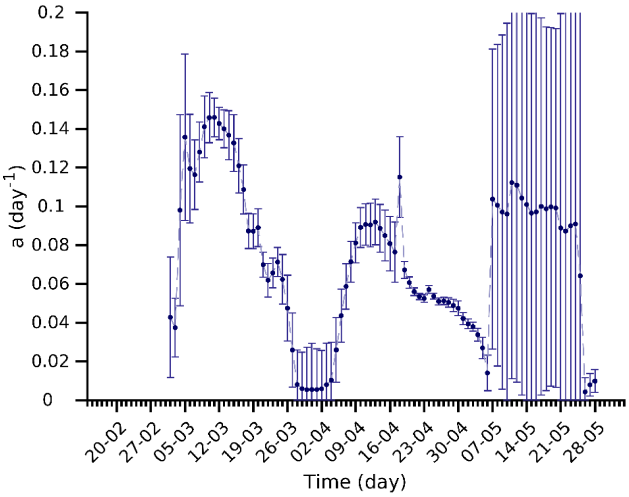


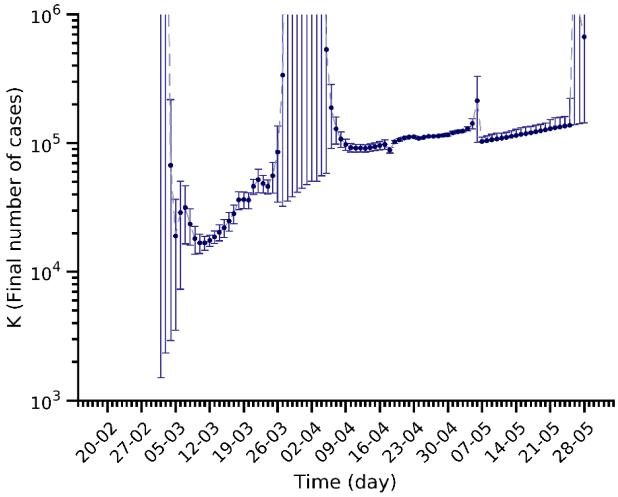


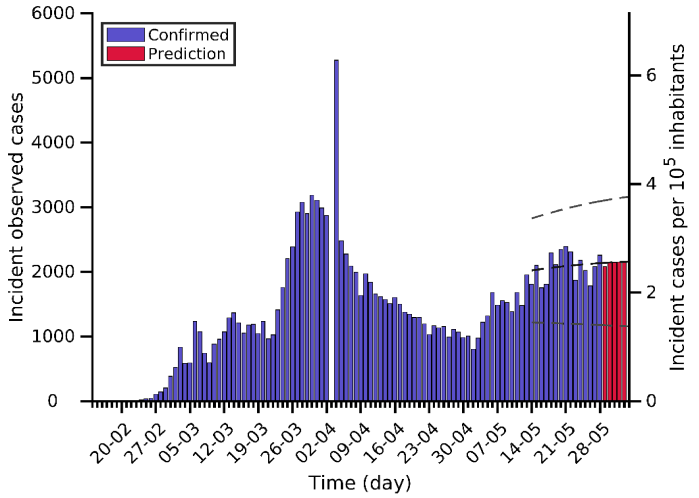
Iran 28-05-2020. Population: 84.0M. Current cumulated incidence: 171/10^5

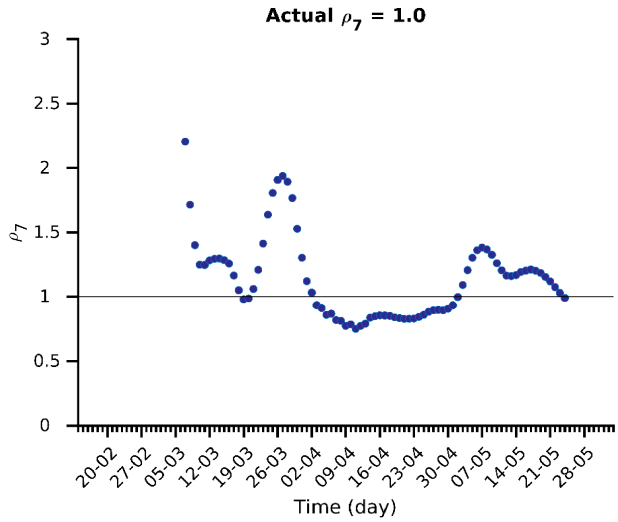


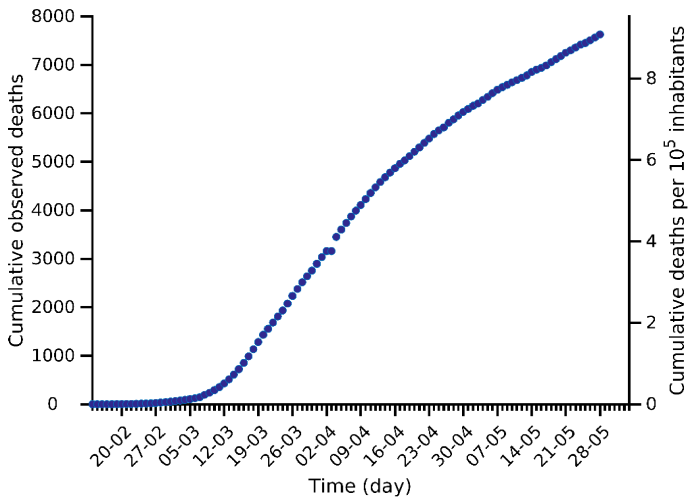


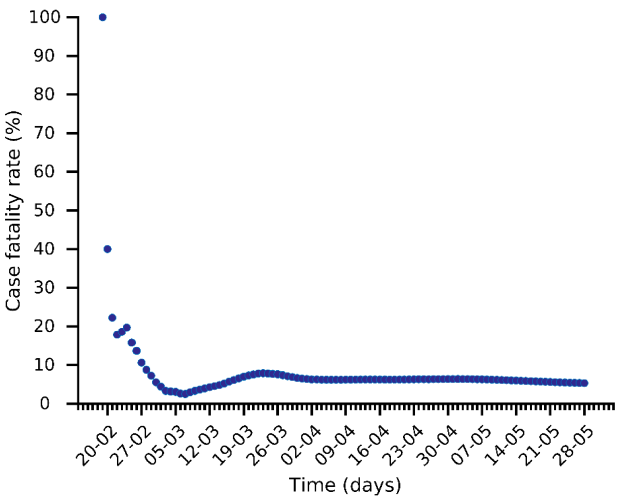




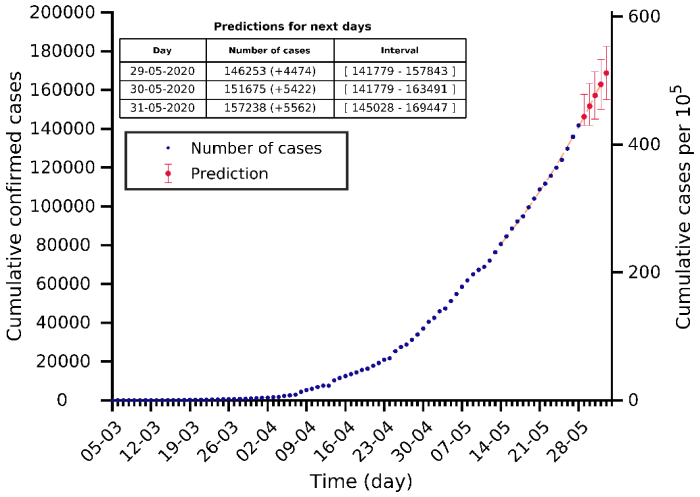


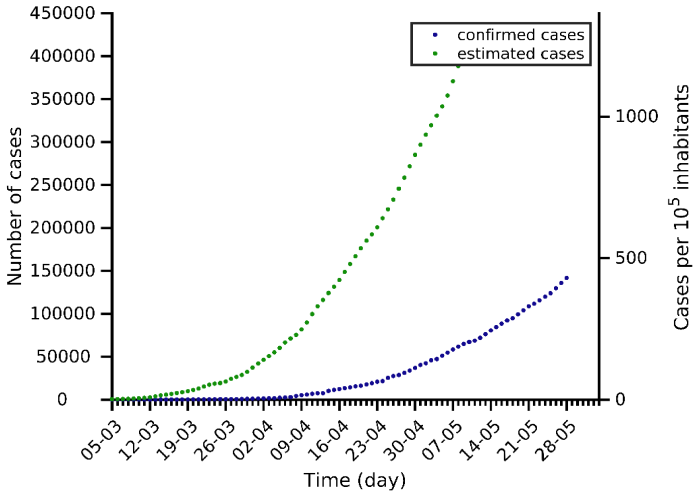


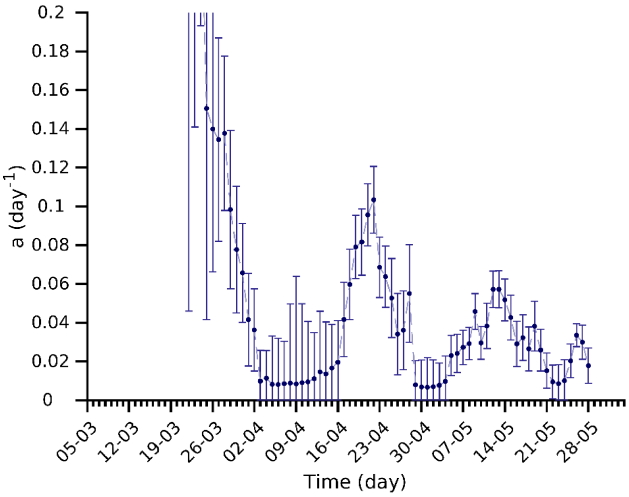


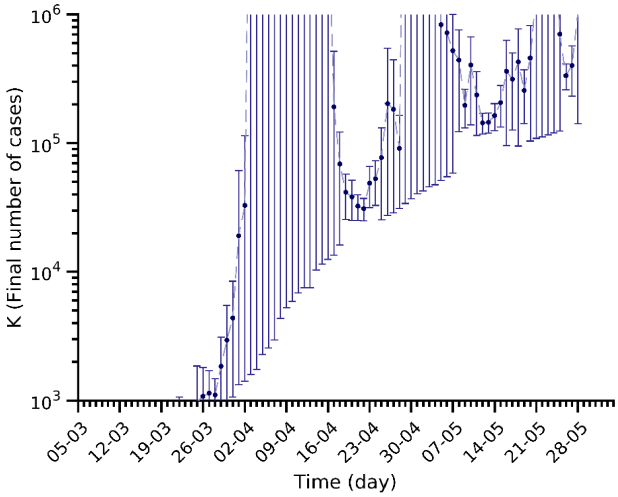


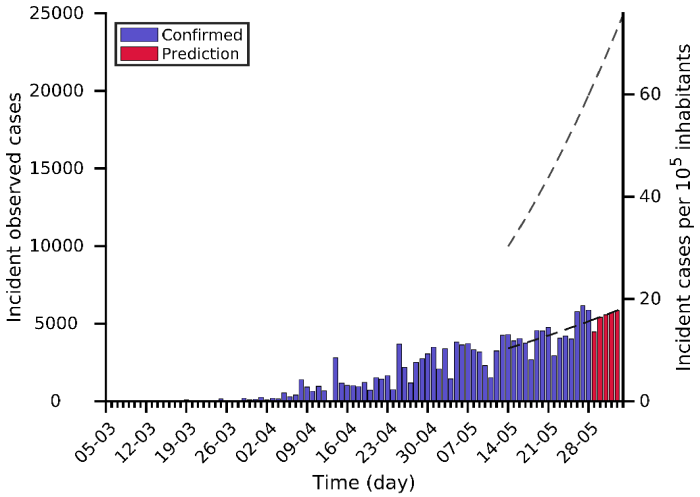
Peru 28-05-2020. Population: 33.0M. Current cumulated incidence: 430/10^5

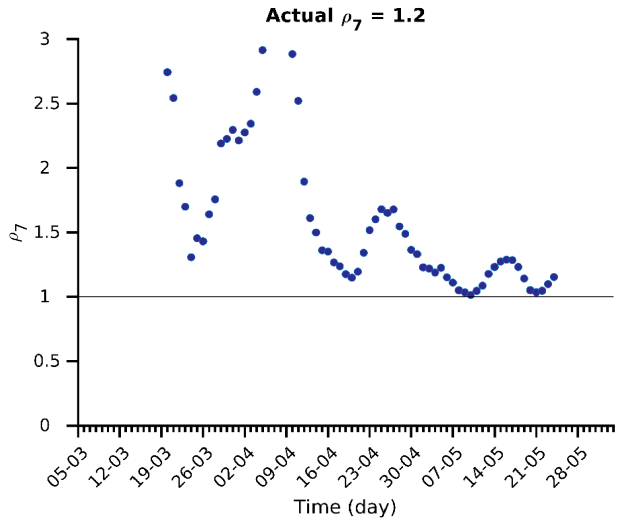


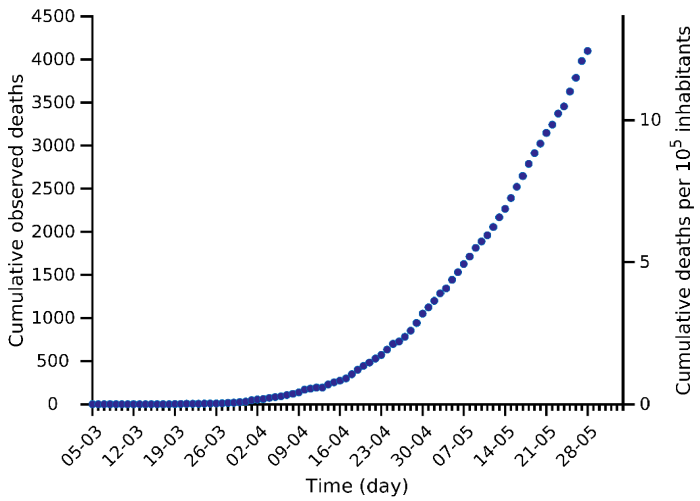


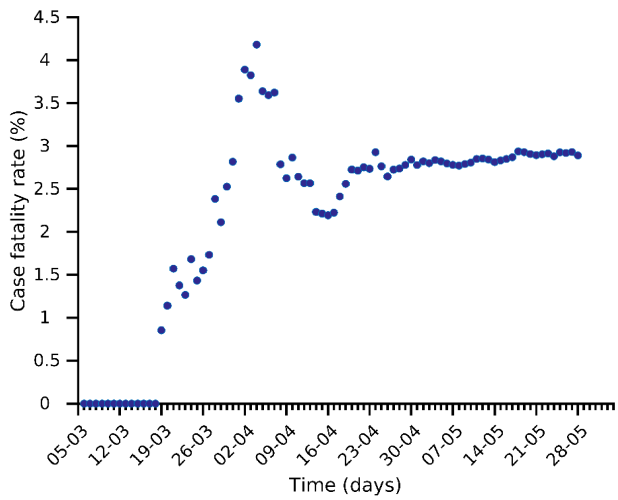




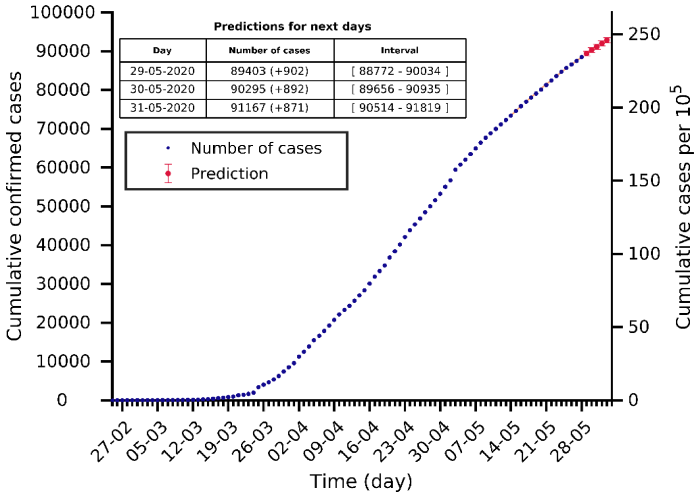


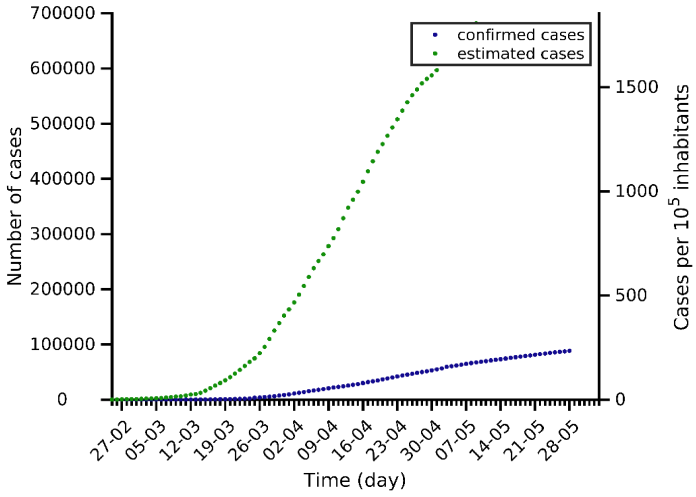


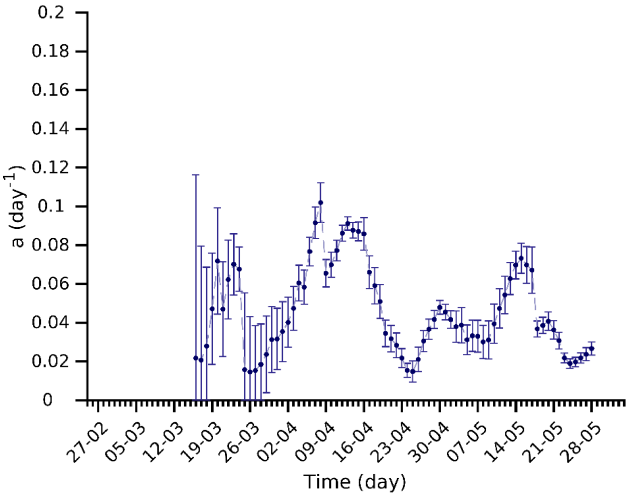


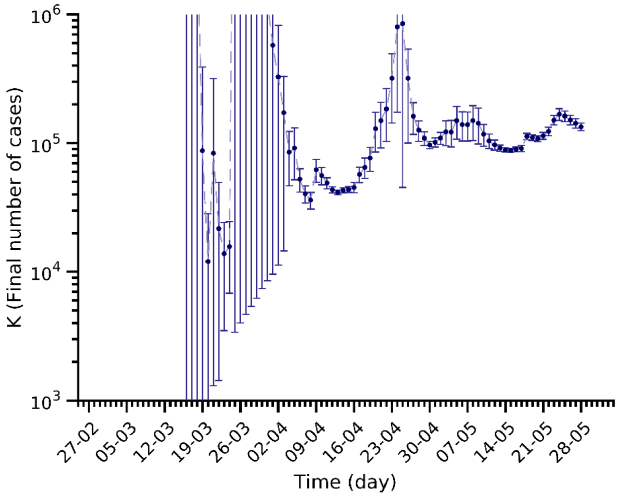


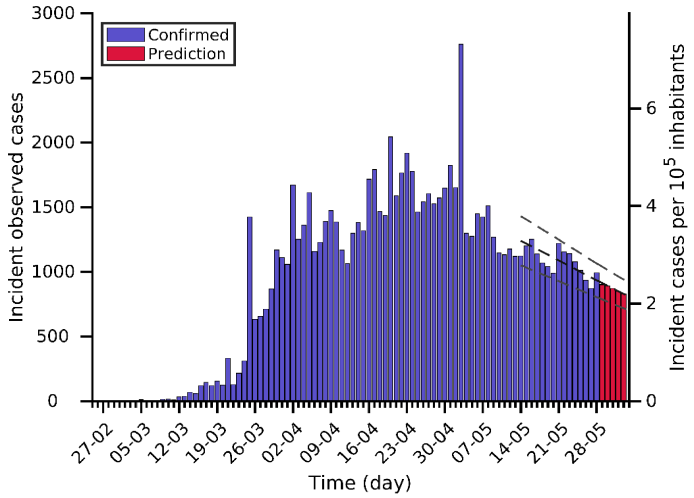
Canada 28-05-2020. Population: 37.7M. Current cumulated incidence: 234/10^5

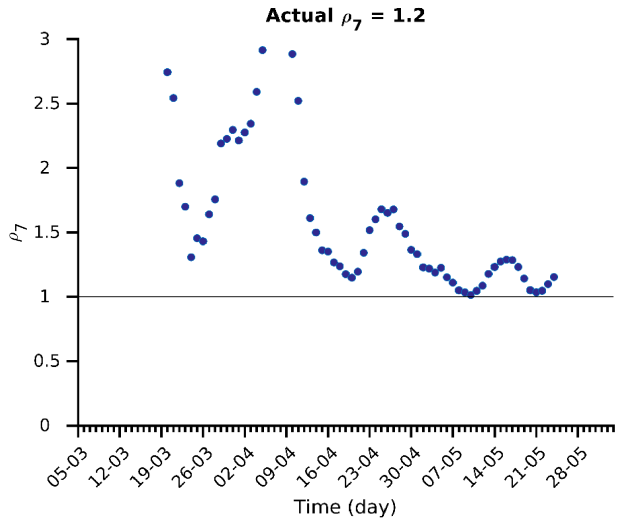


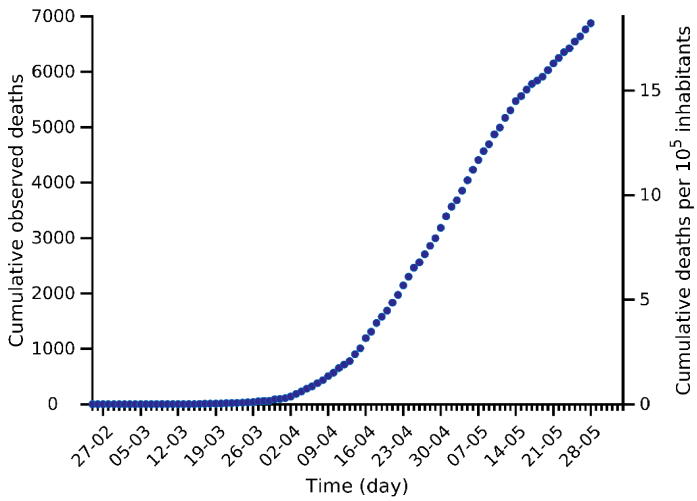


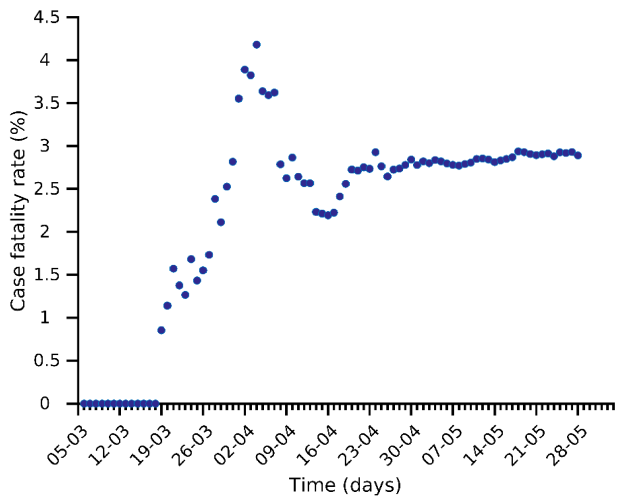




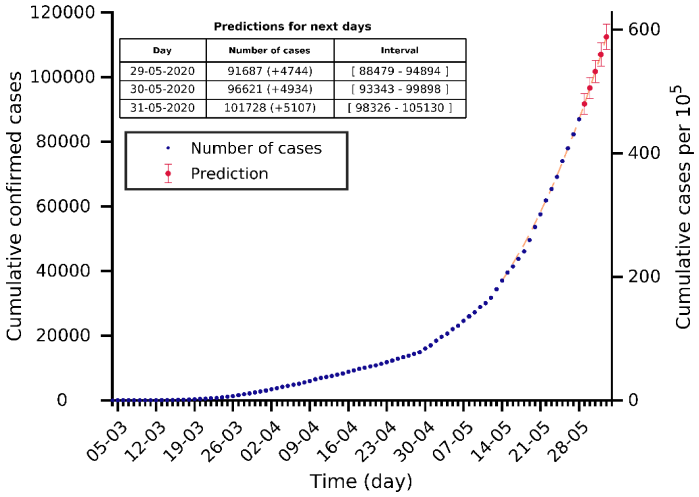


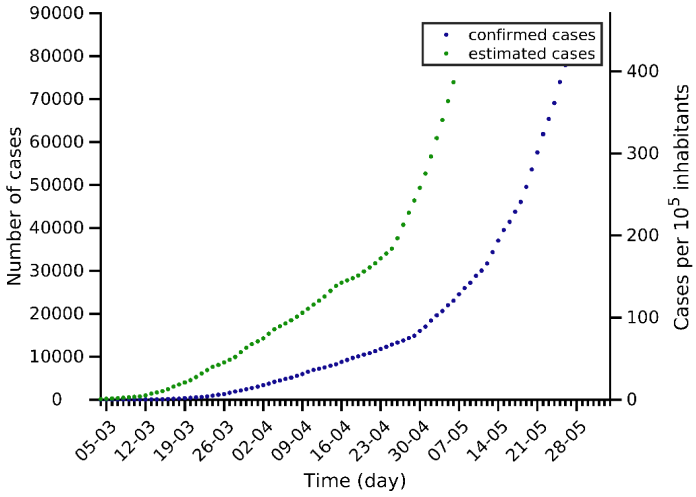


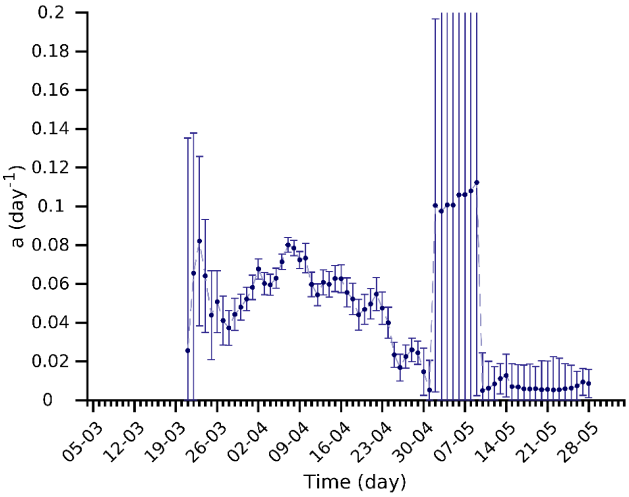


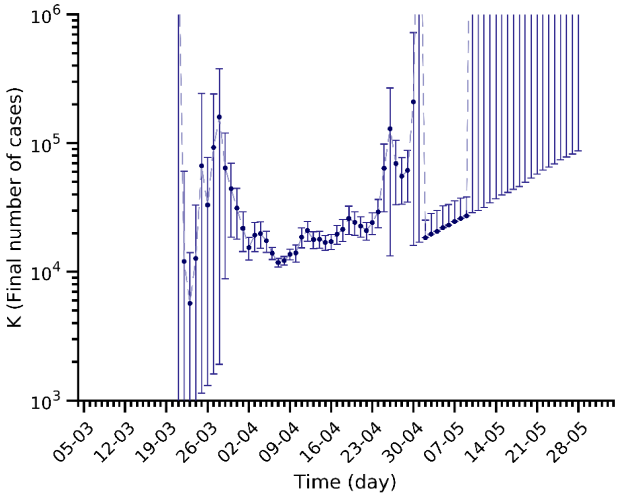


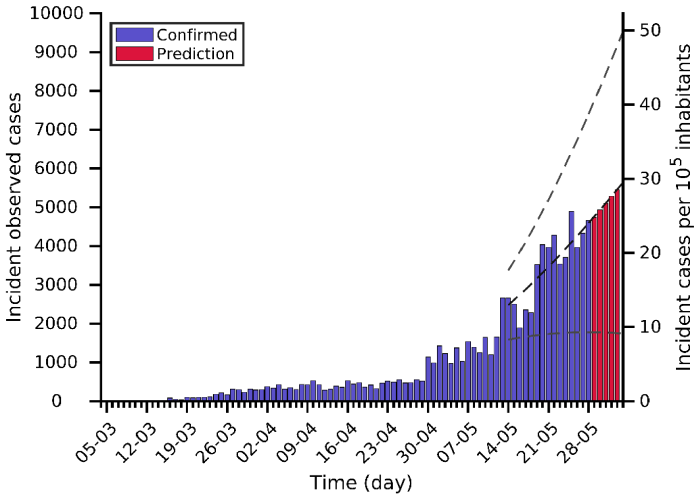
Chile 28-05-2020. Population: 19.1M. Current cumulated incidence: 455/10^5

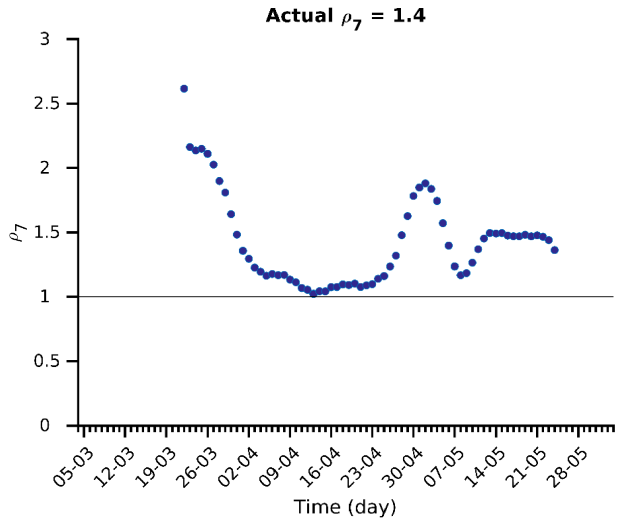


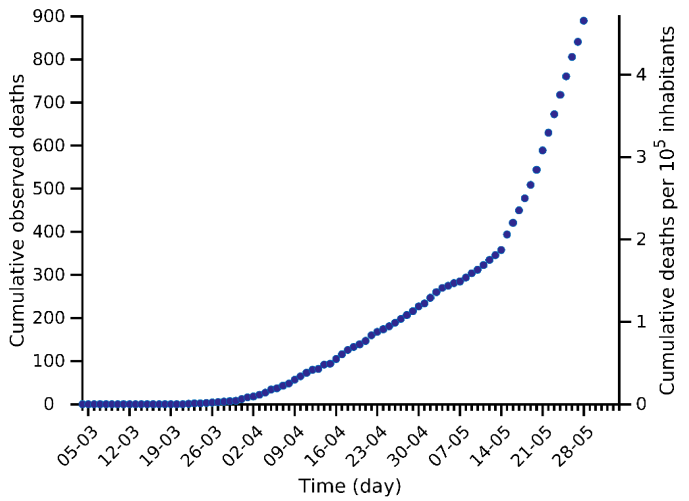


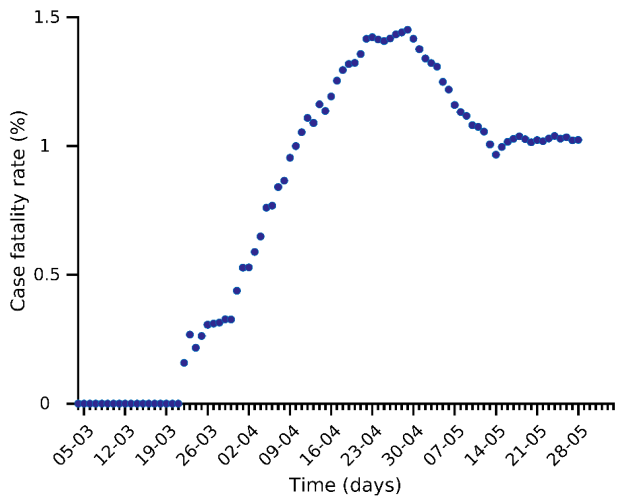




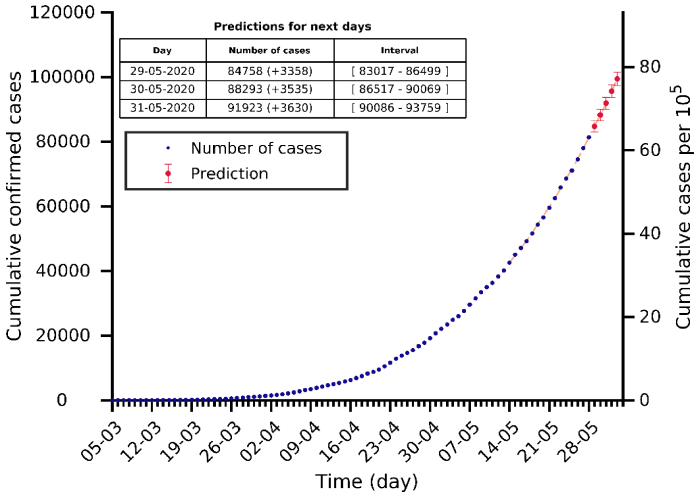


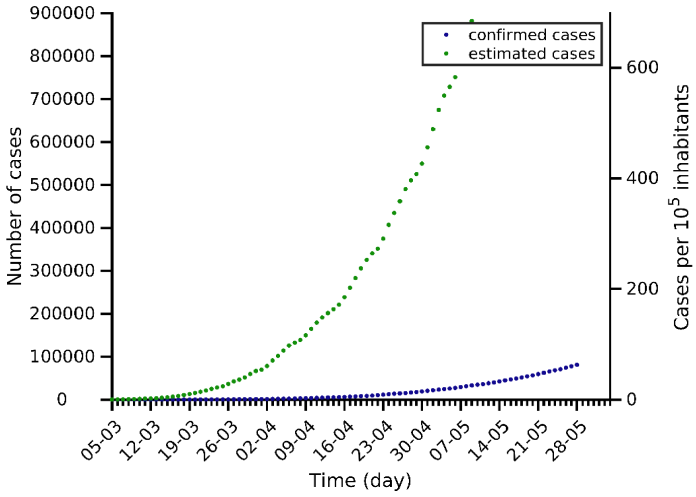


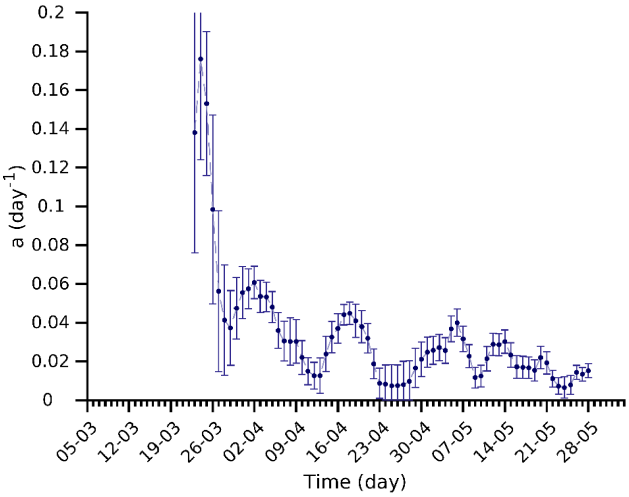


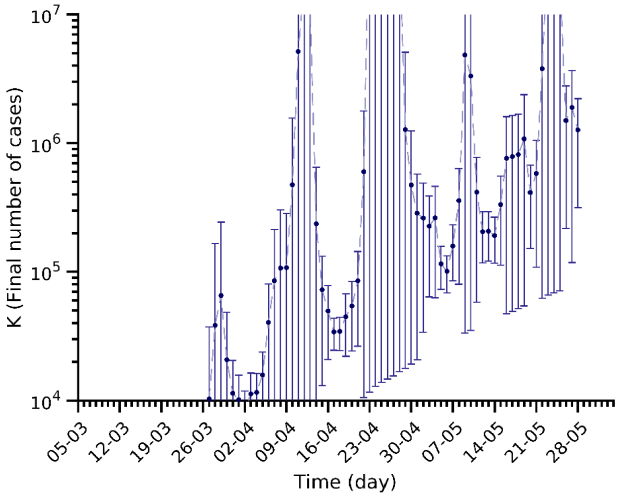


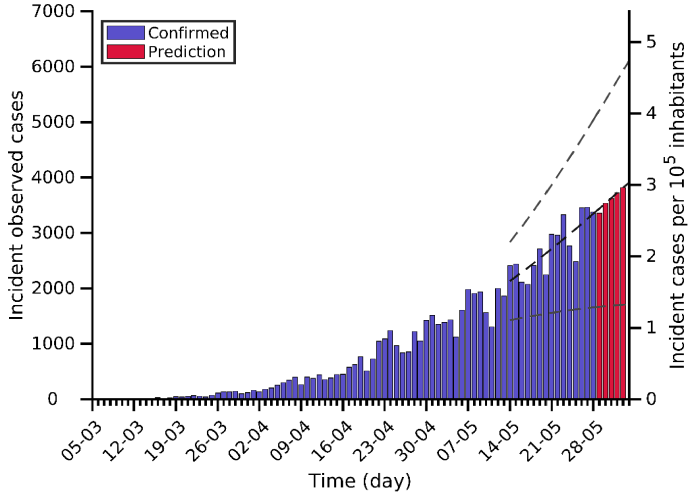
Mexico 28-05-2020. Population: 128.9M. Current cumulated incidence: $63/10^5$

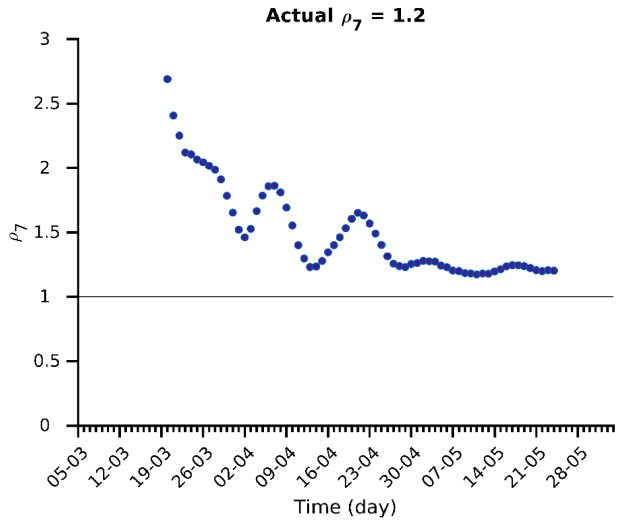


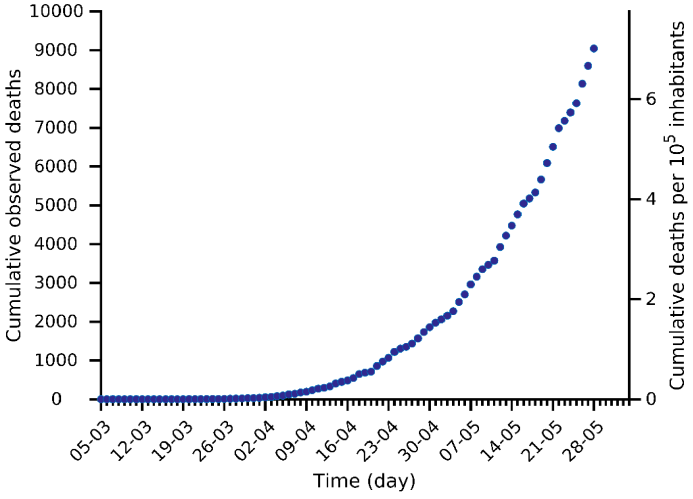


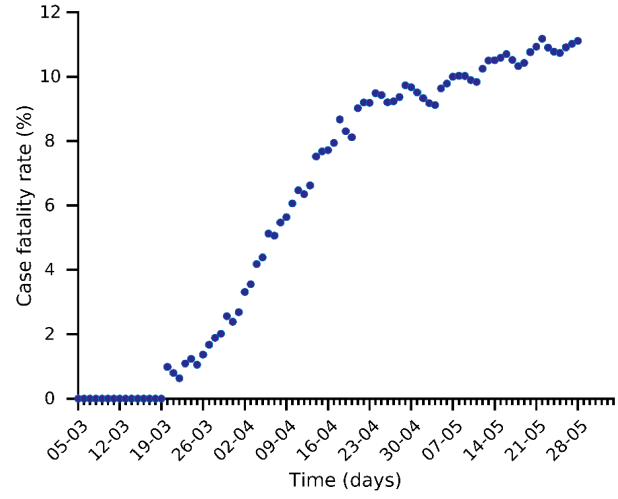




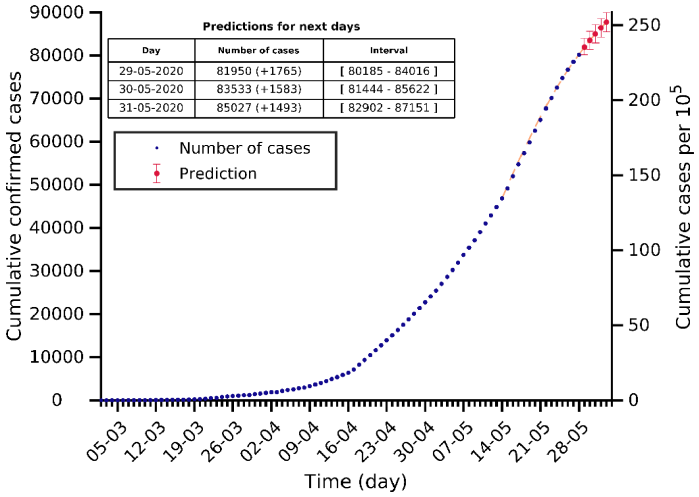


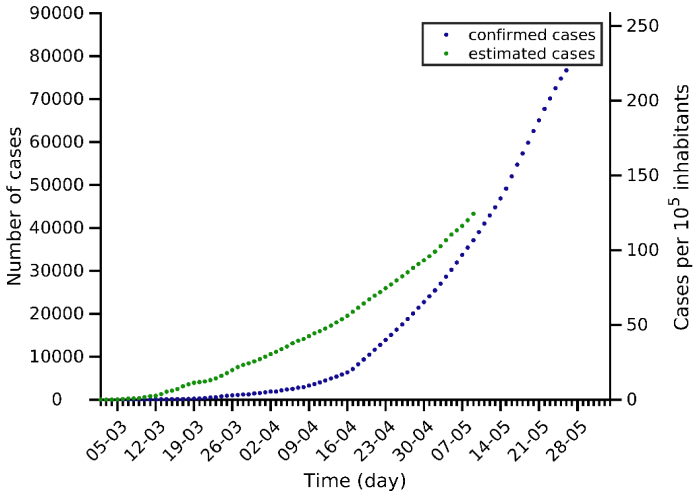


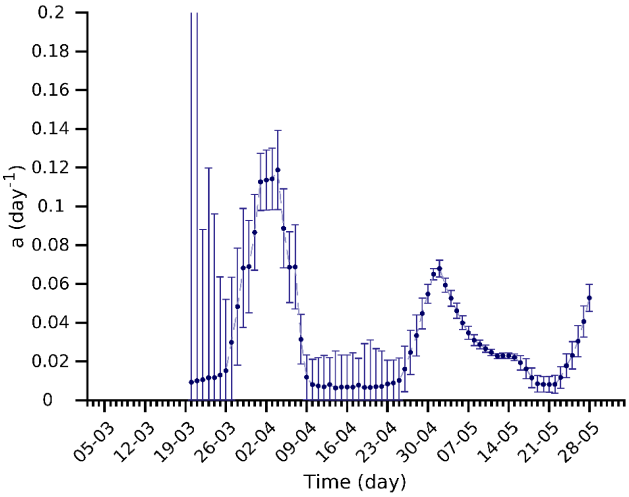


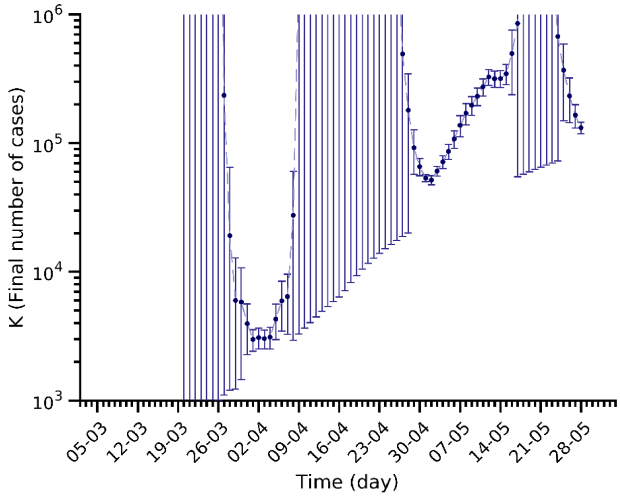


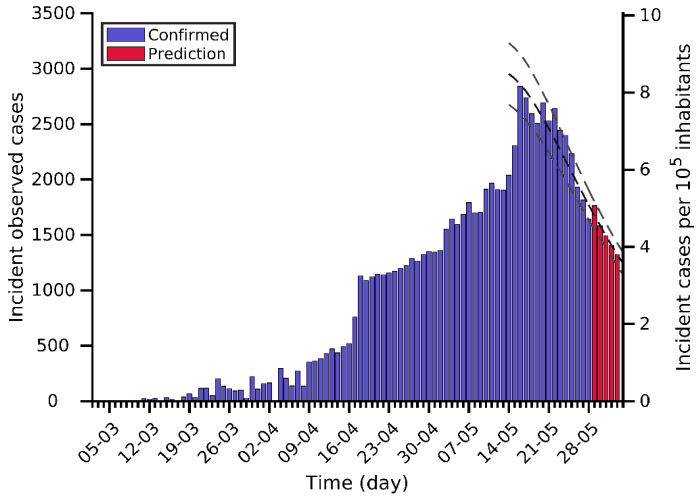
Saudi Arabia 28-05-2020. Population: 34.8M. Current cumulated incidence: 230/10^5

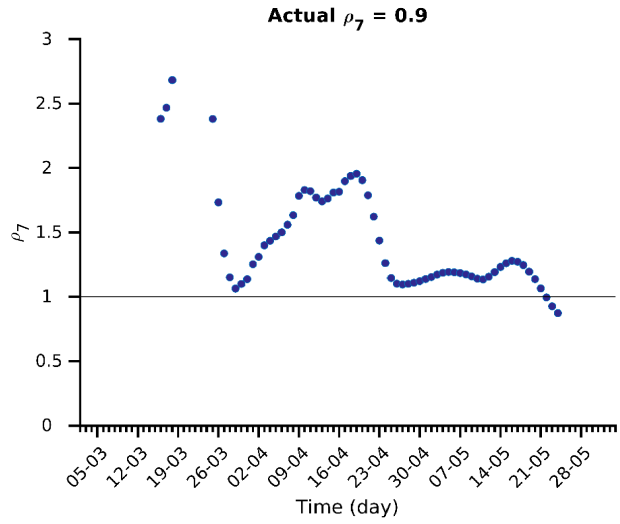


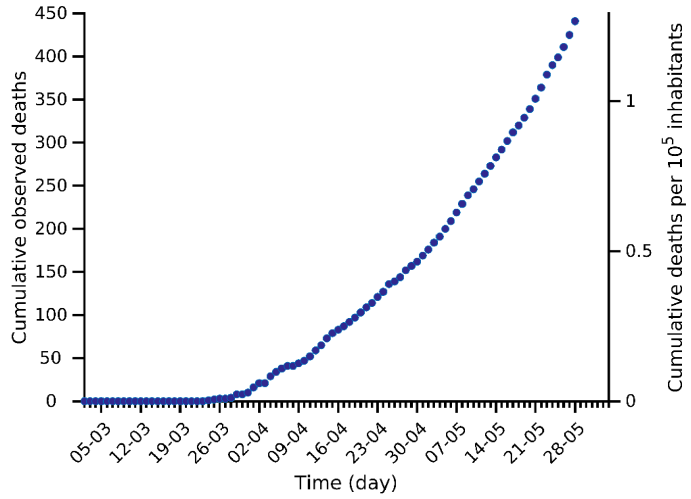


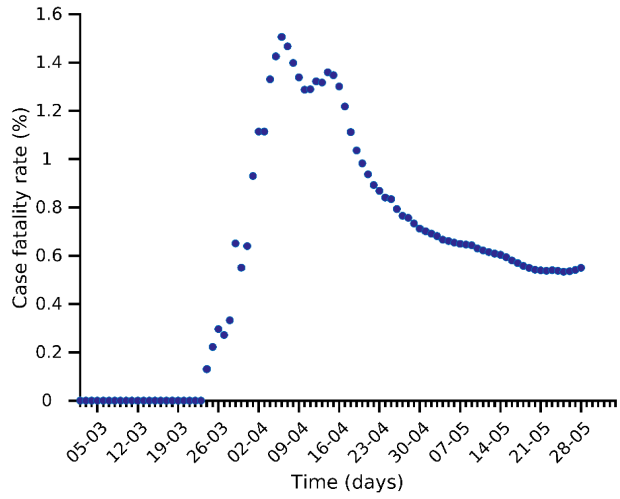




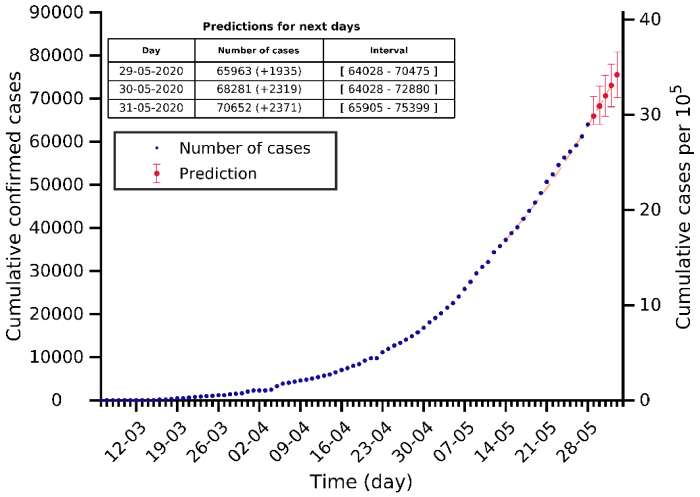


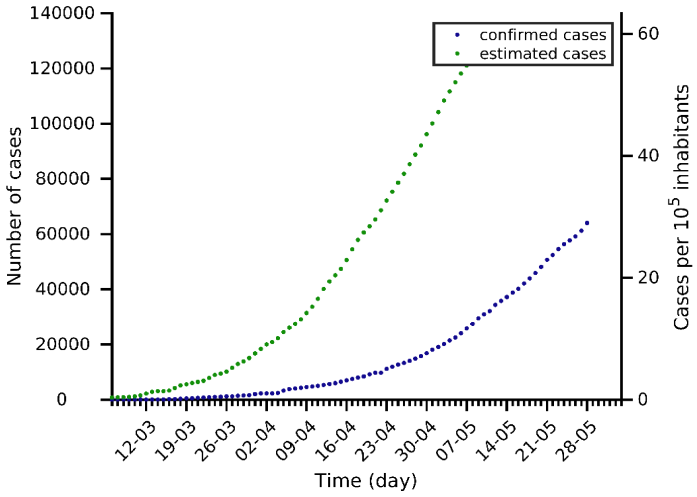


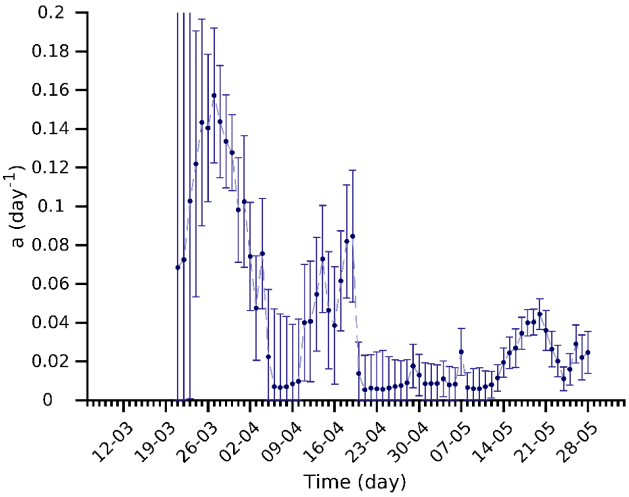


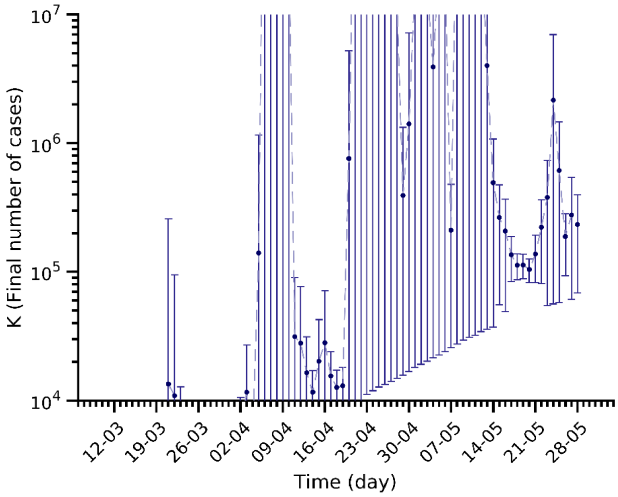


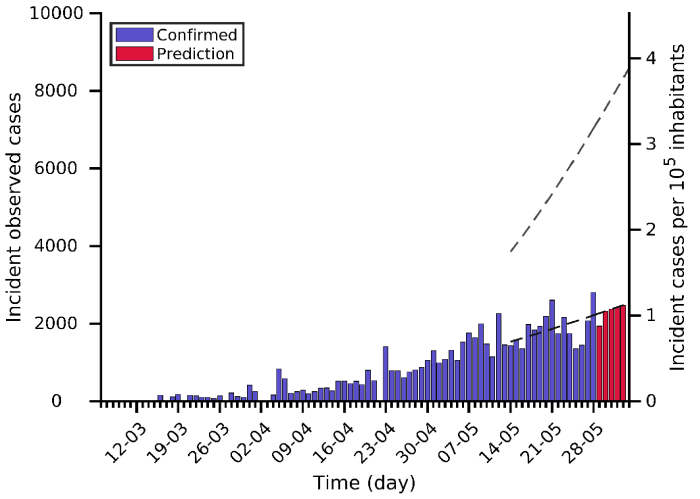
Pakistan 28-05-2020. Population: 220.9M. Current cumulated incidence: $29/10^5$

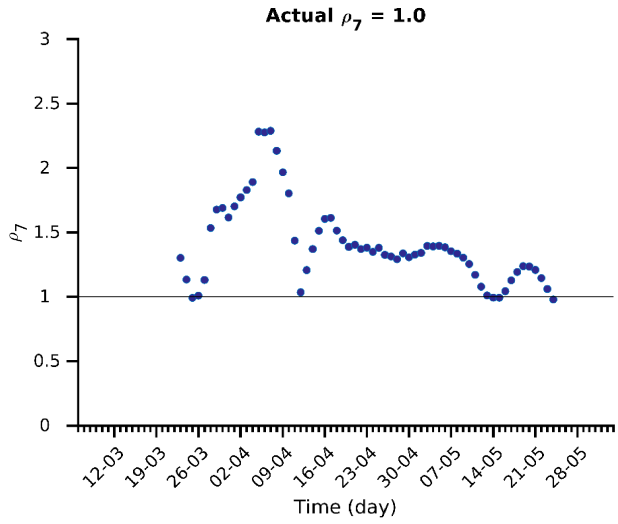


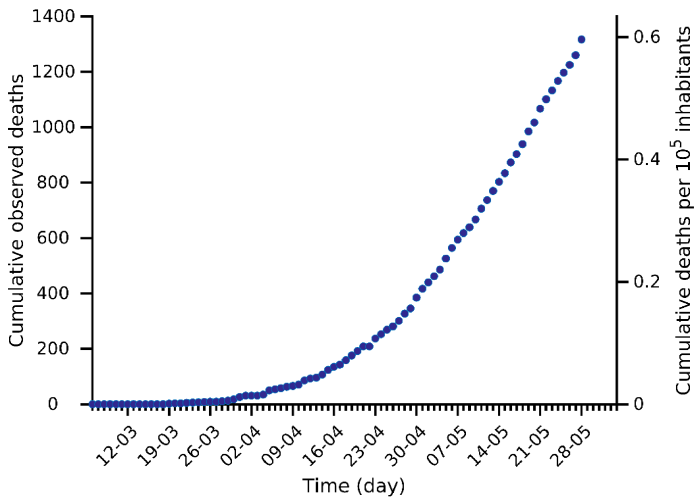


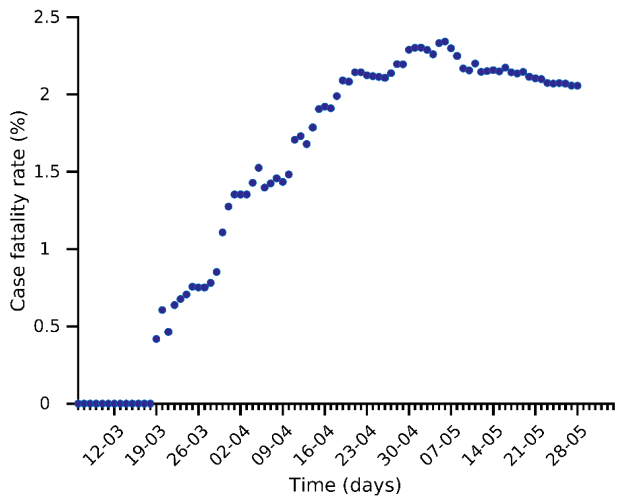




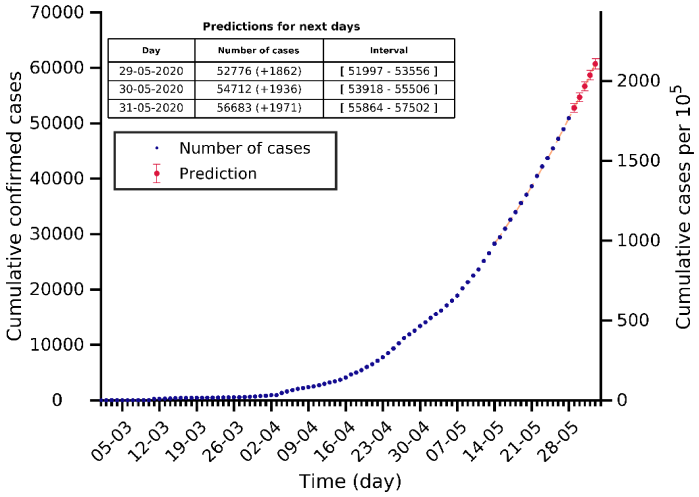


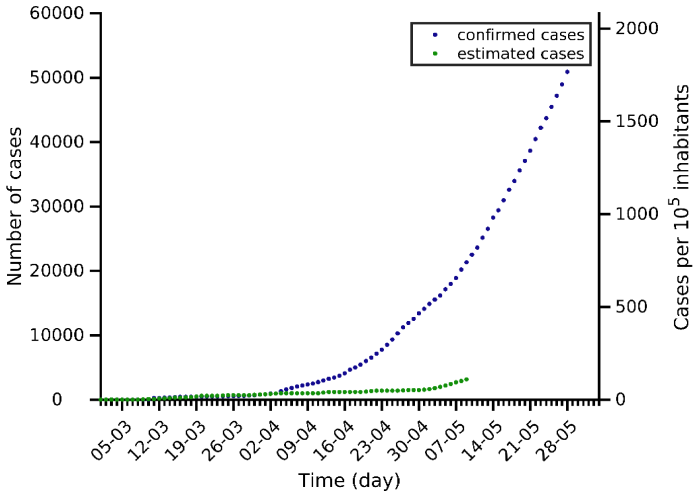


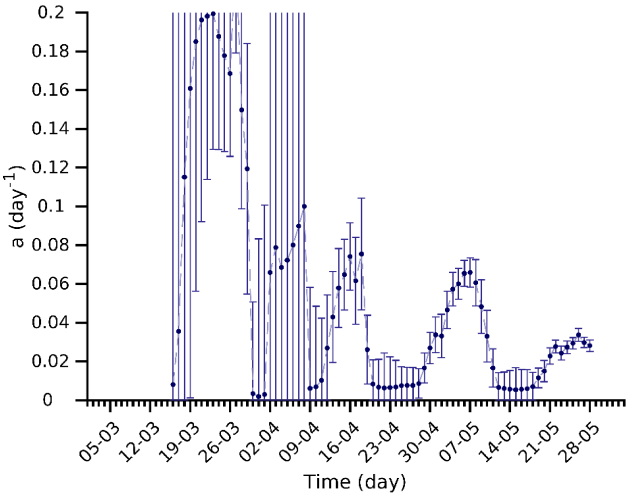


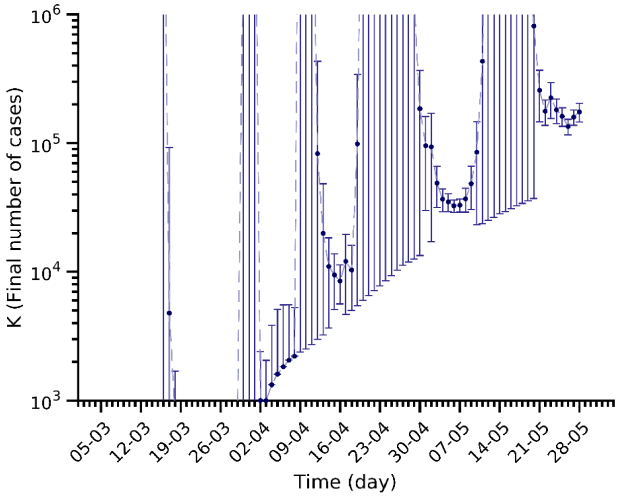


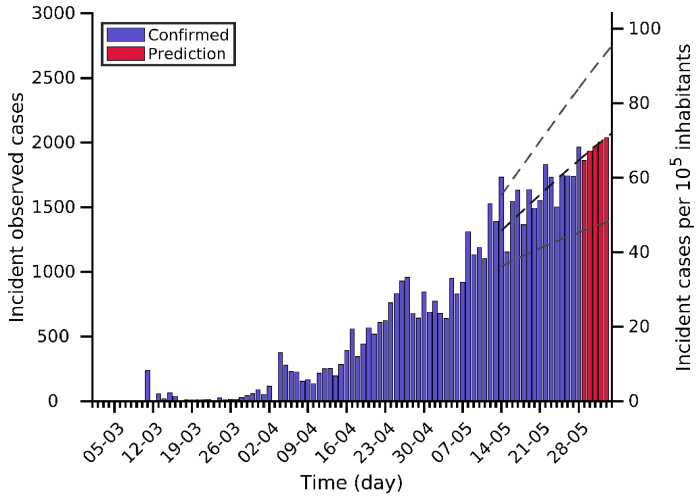
Qatar 28-05-2020. Population: 2.9M. Current cumulated incidence: 1767/10^5

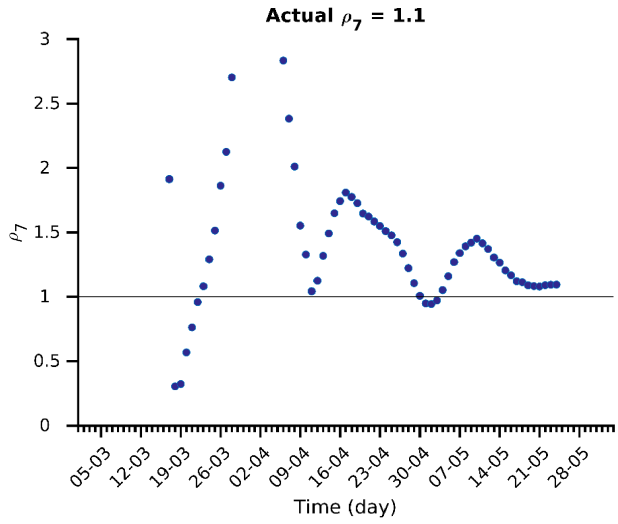


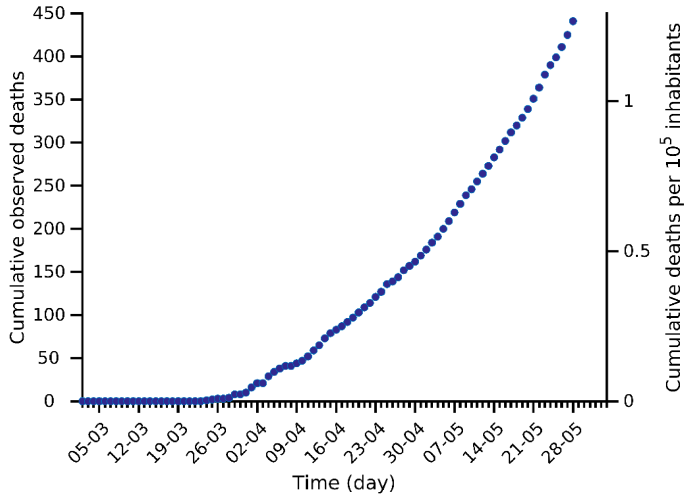


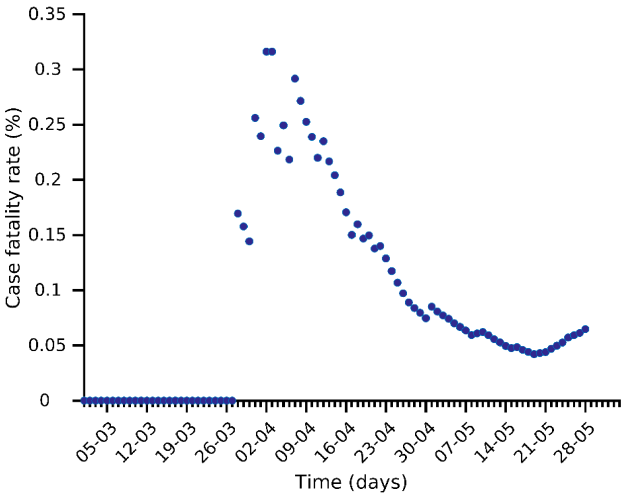




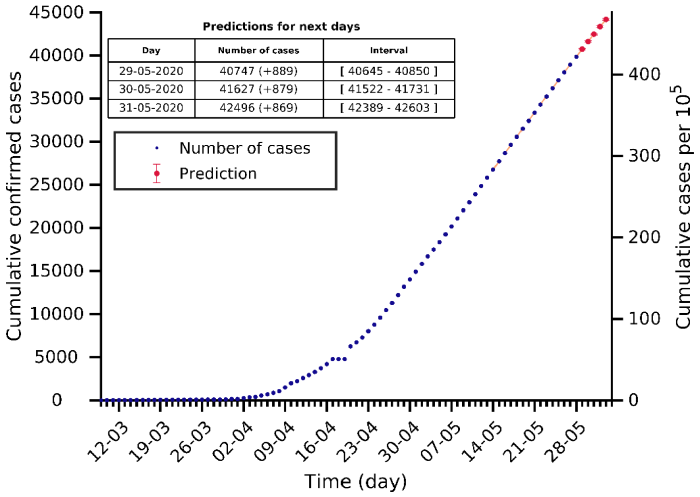


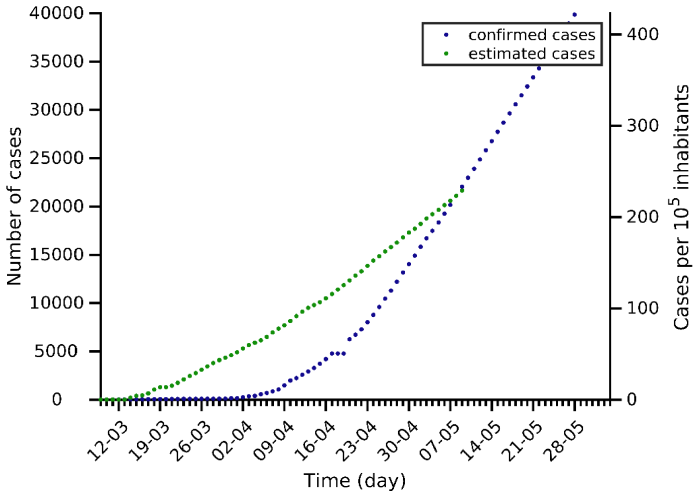


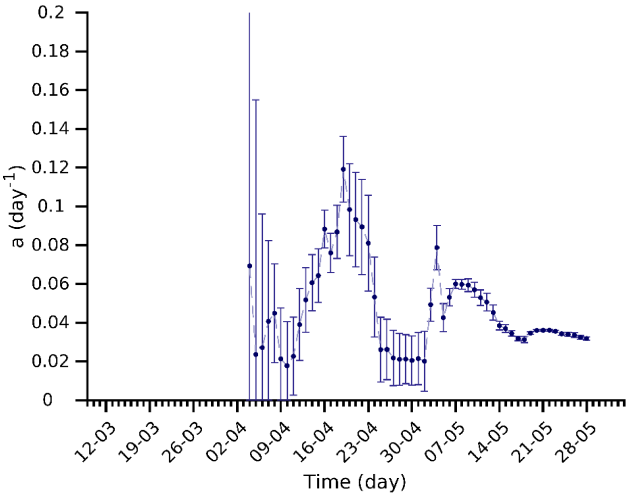


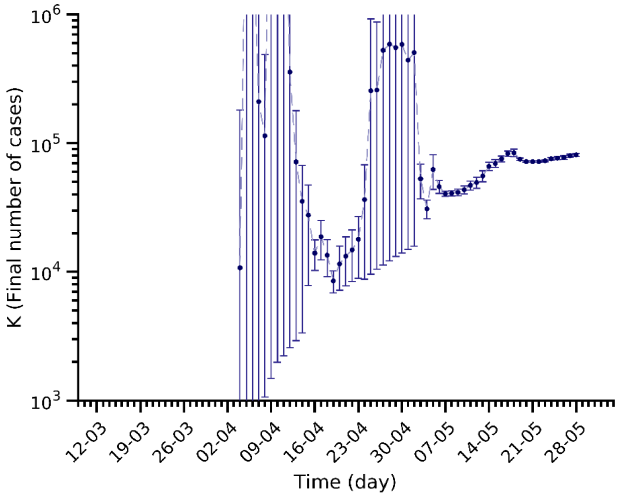


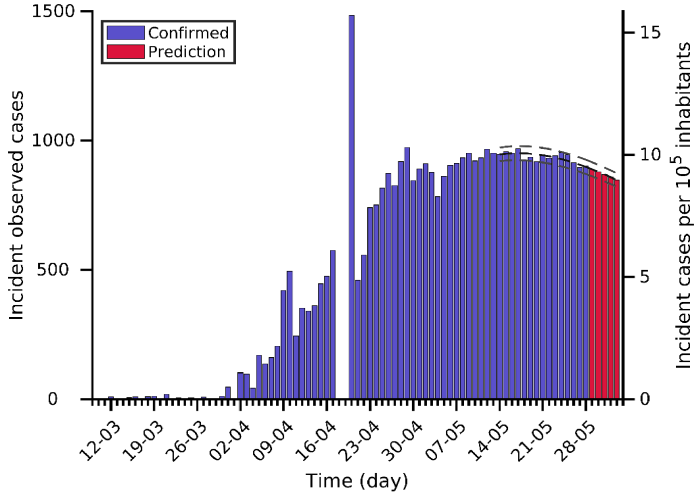
Belarus 28-05-2020. Population: 9.4M. Current cumulated incidence: 422/10^5

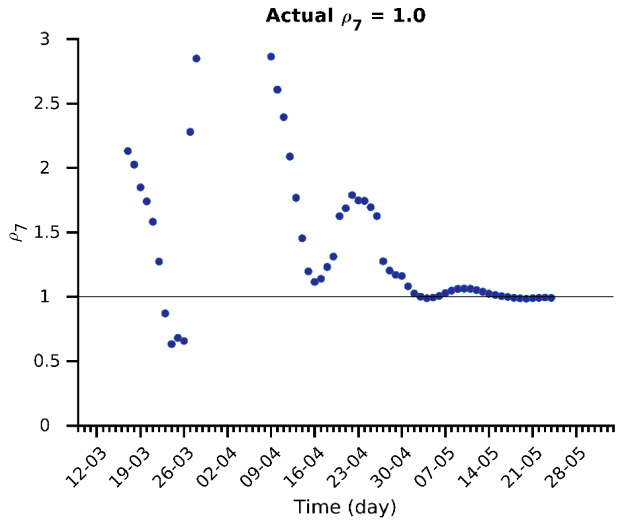


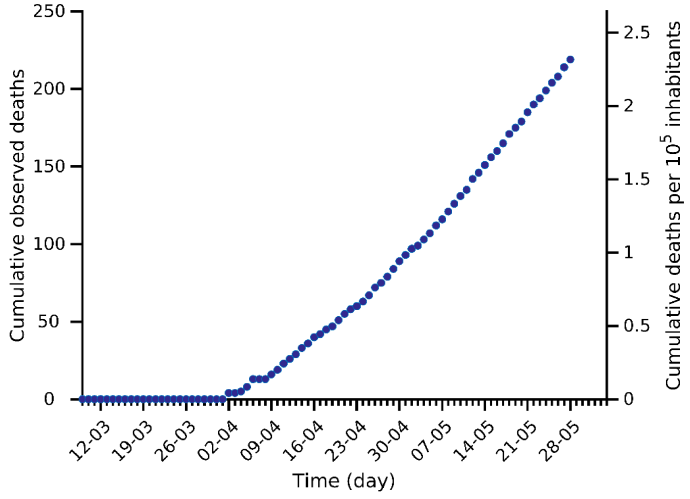


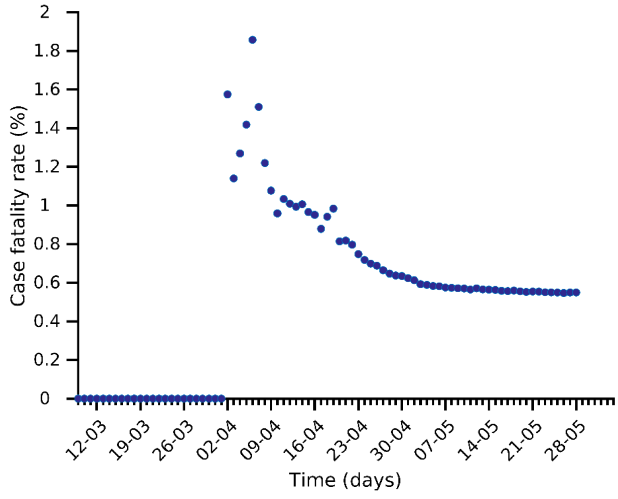




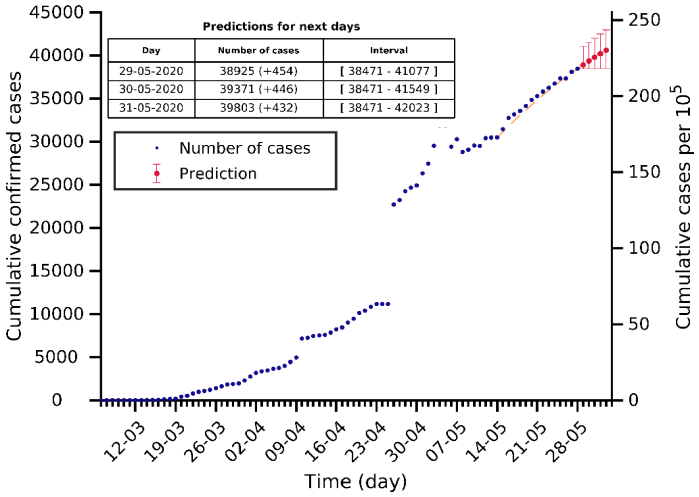


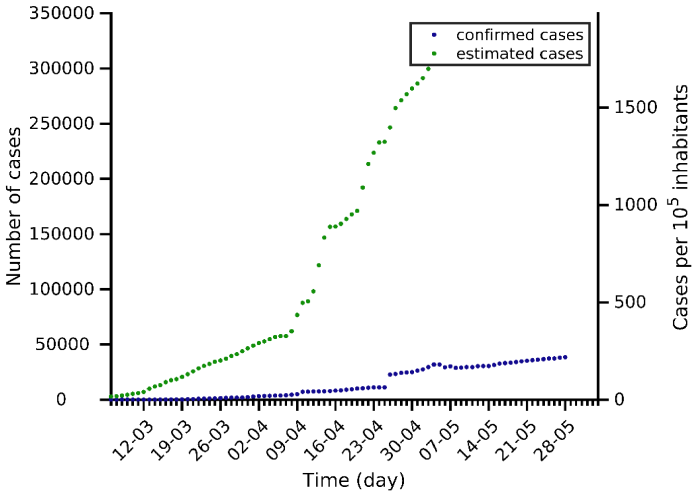


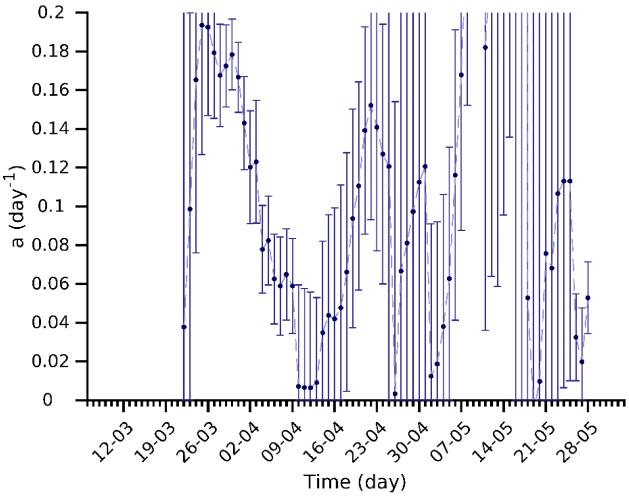


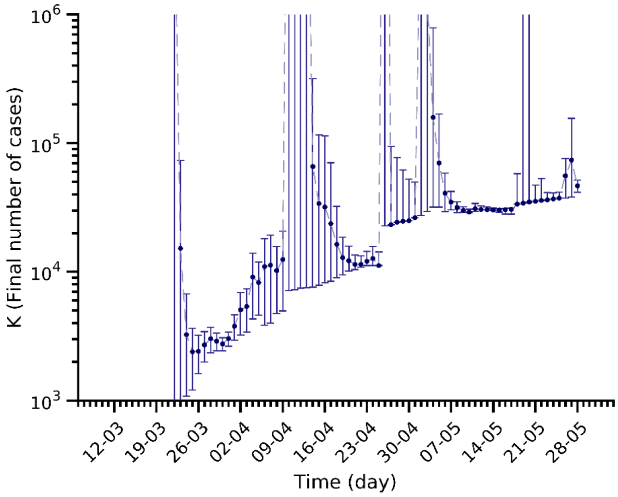


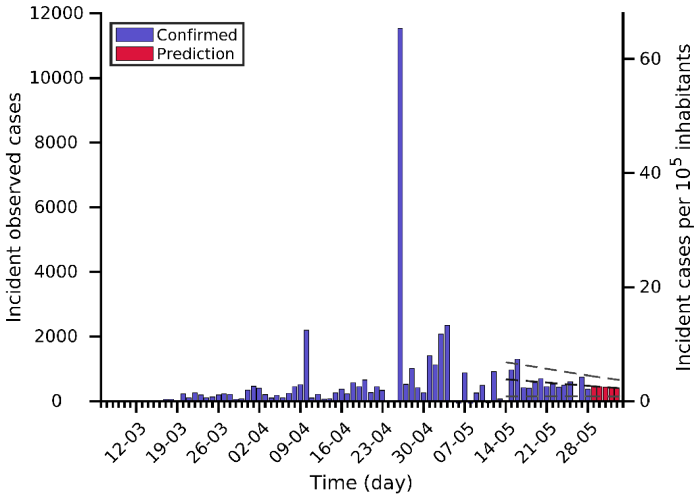
Ecuador 28-05-2020. Population: 17.6M. Current cumulated incidence: 218/10^5

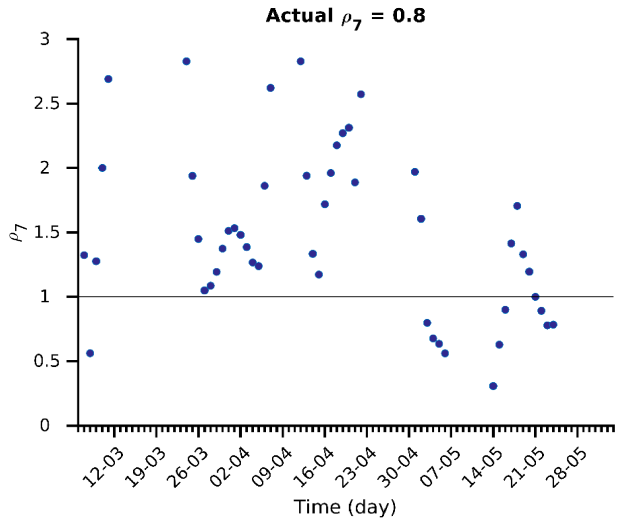


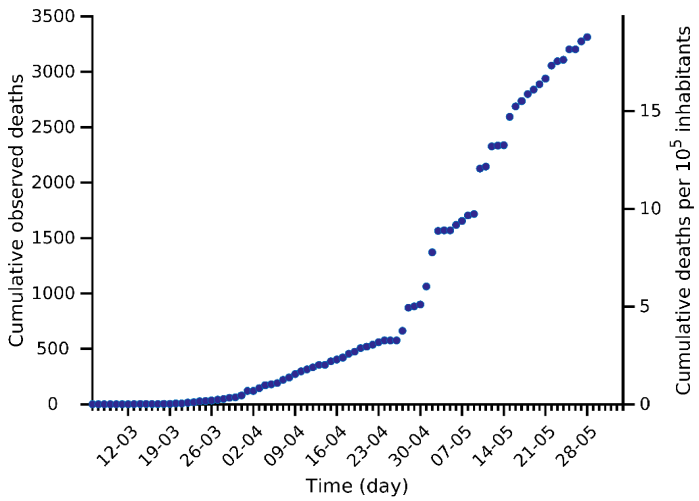


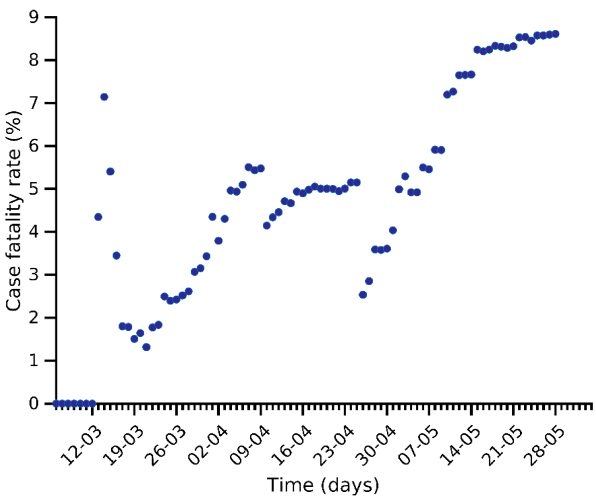




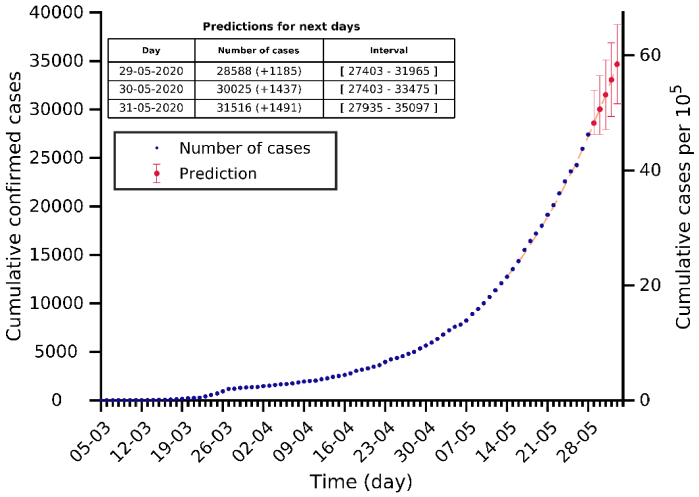


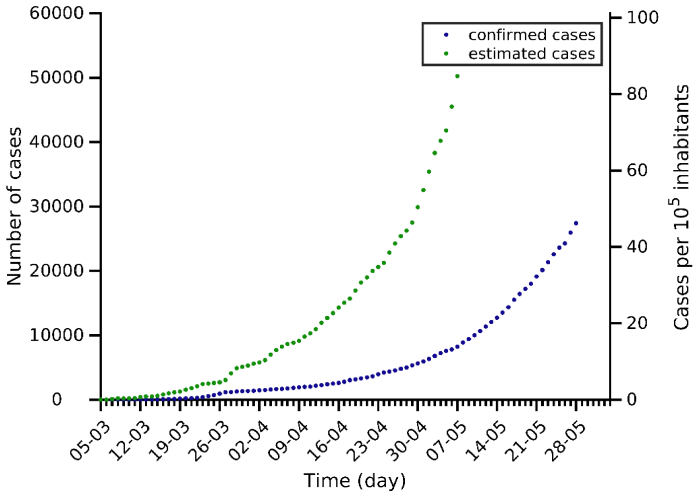


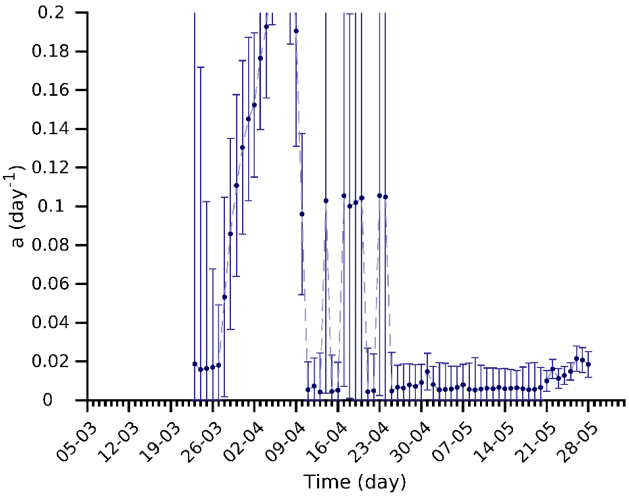


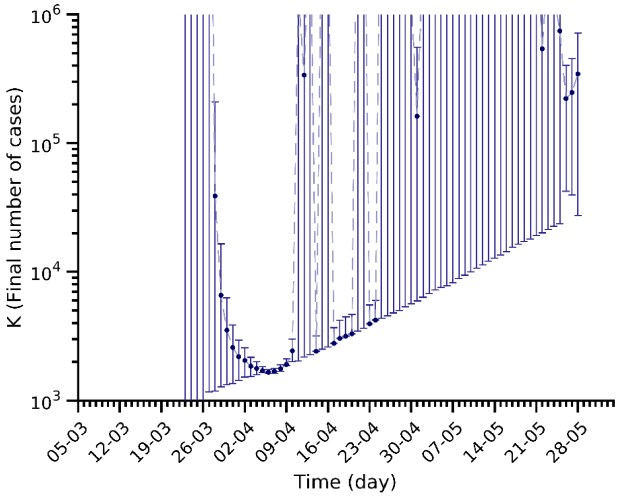


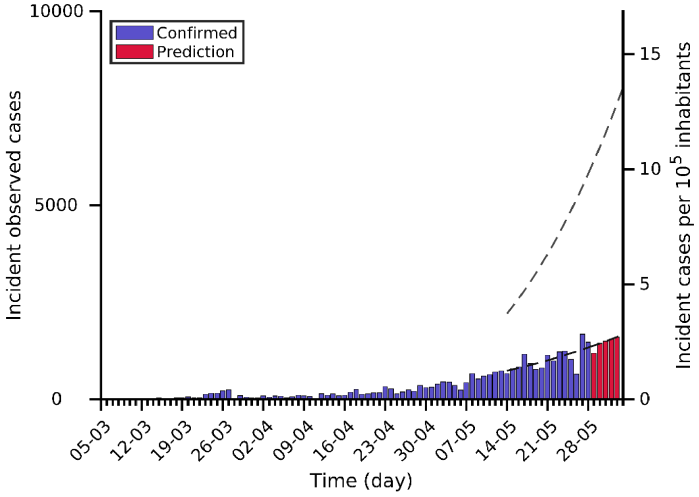
South Africa 28-05-2020. Population: 59.3M. Current cumulated incidence: 46/10^5

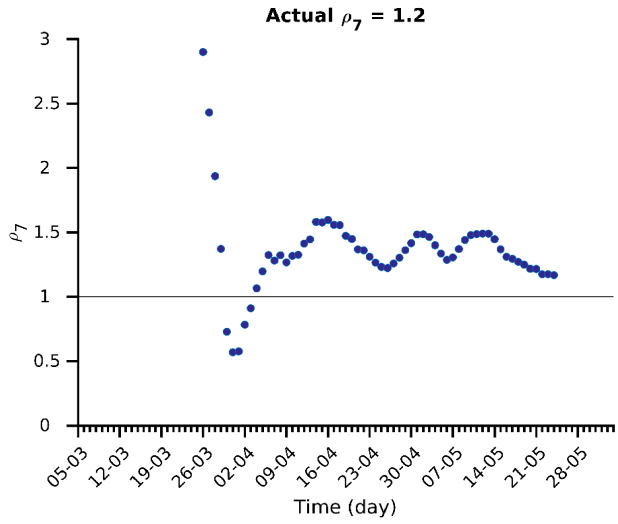


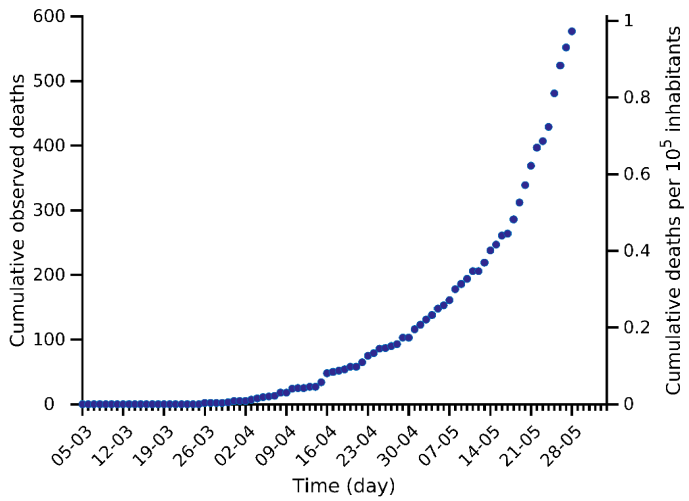


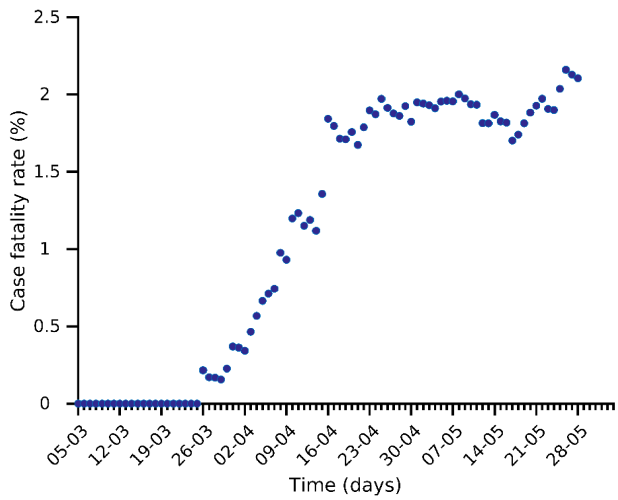




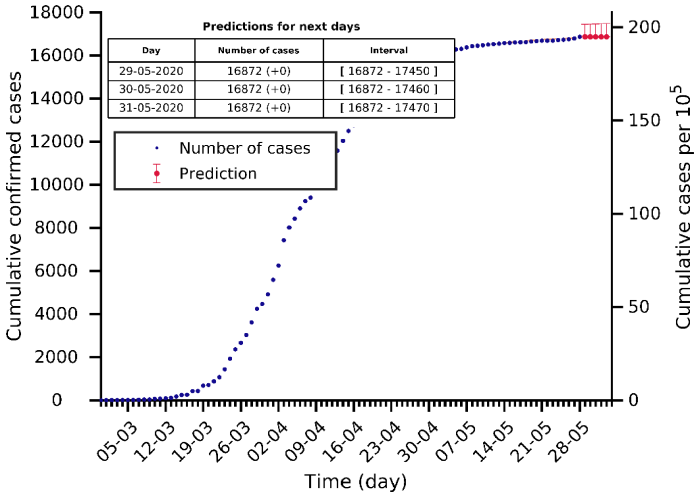


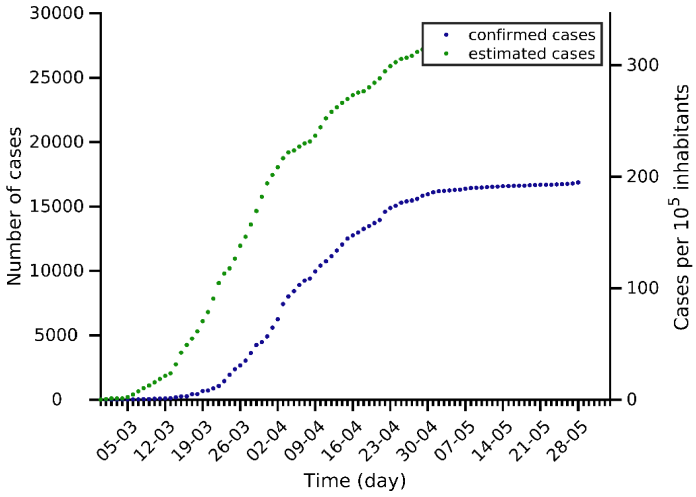


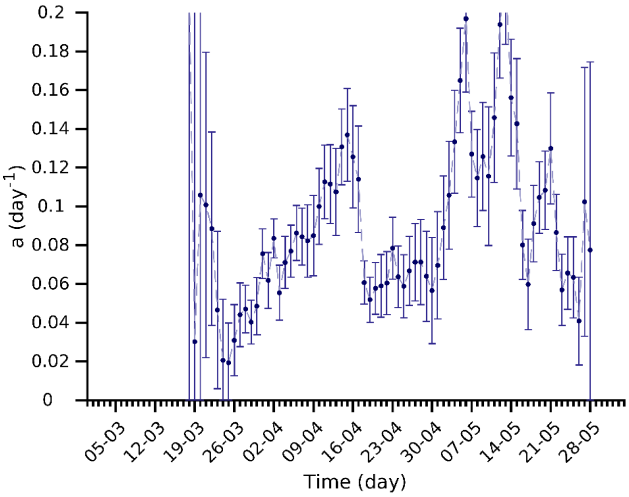


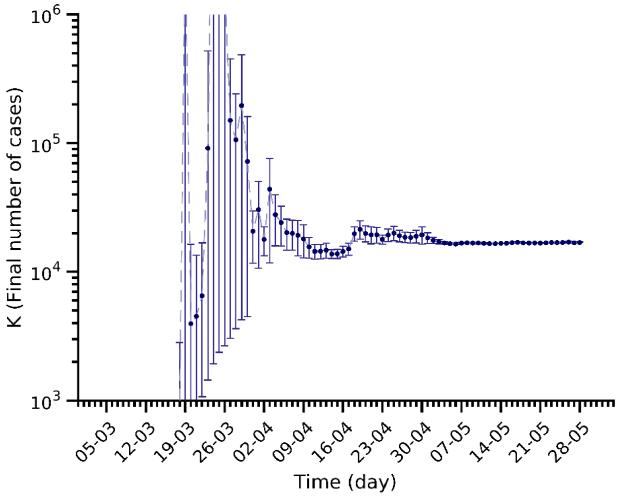


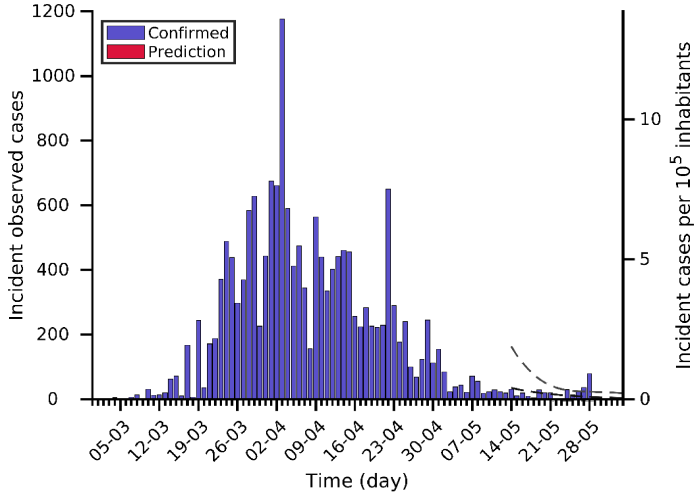
Israel 28-05-2020. Population: 8.7M. Current cumulated incidence: 195/10⁵

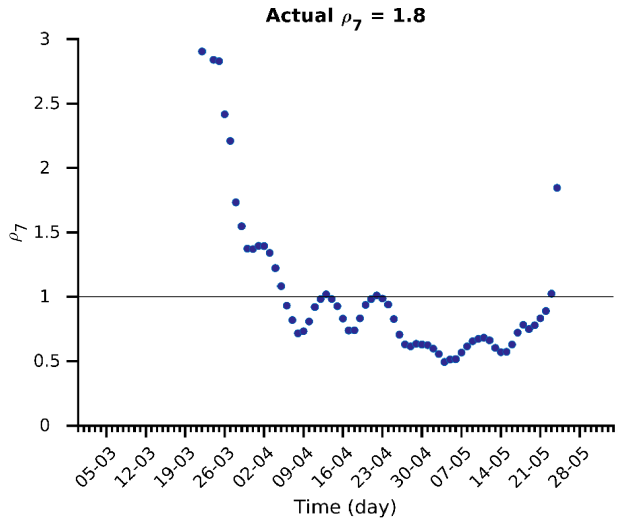


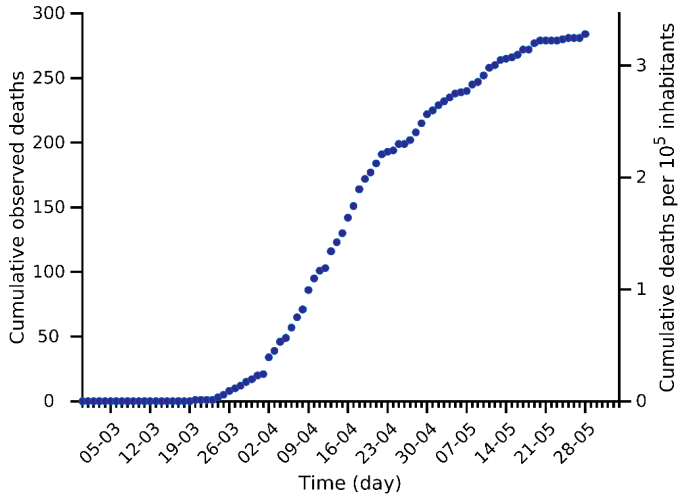


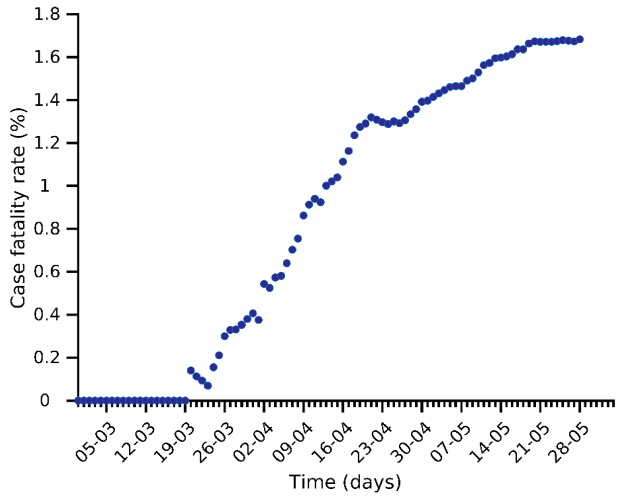




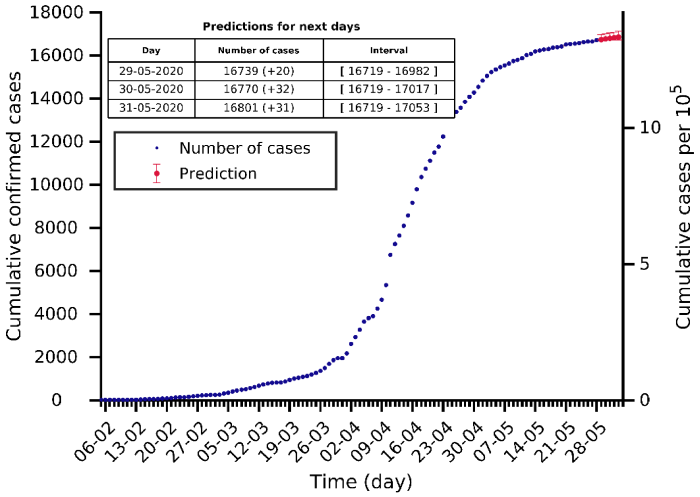


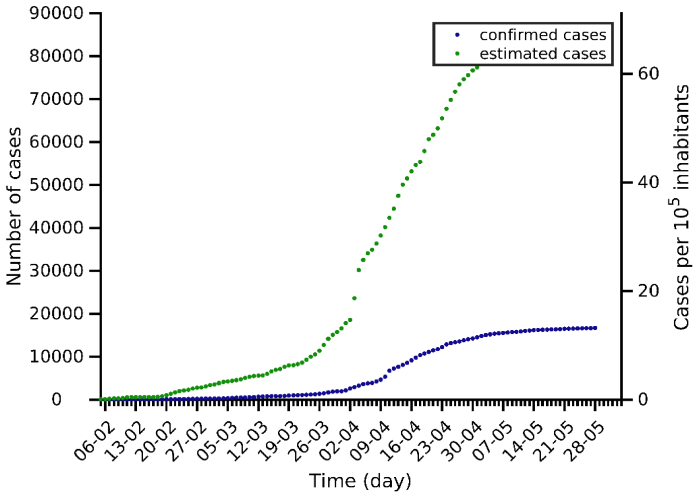


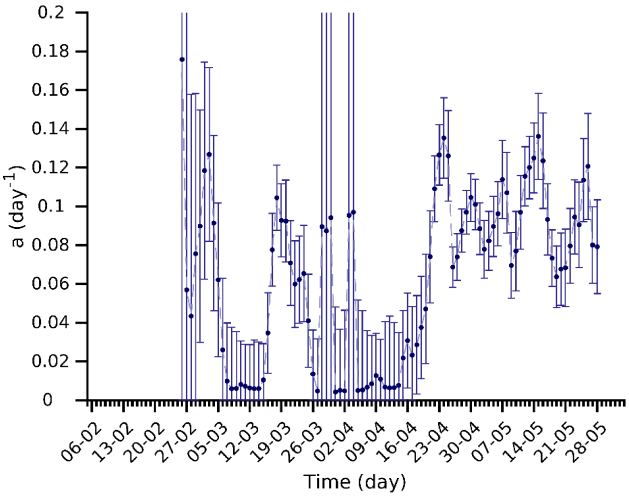


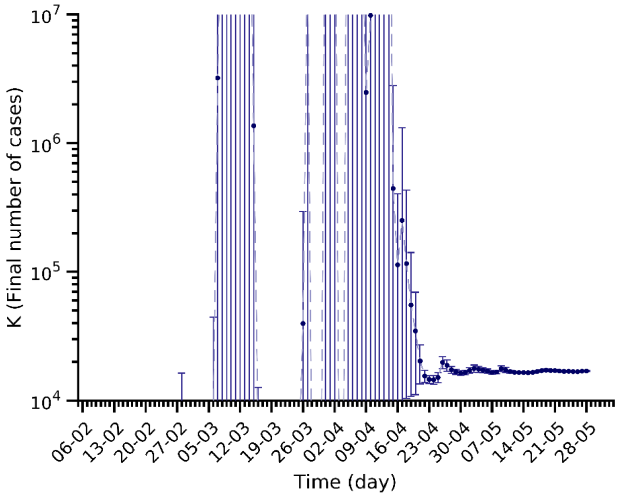


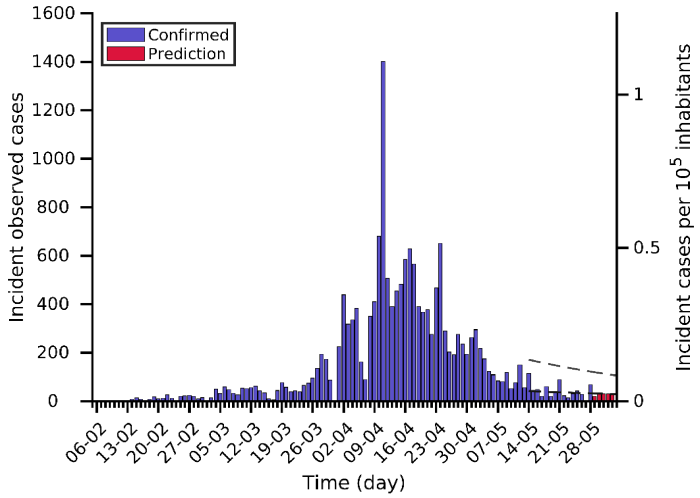
Japan 28-05-2020. Population: 126.5M. Current cumulated incidence: 13/10^5

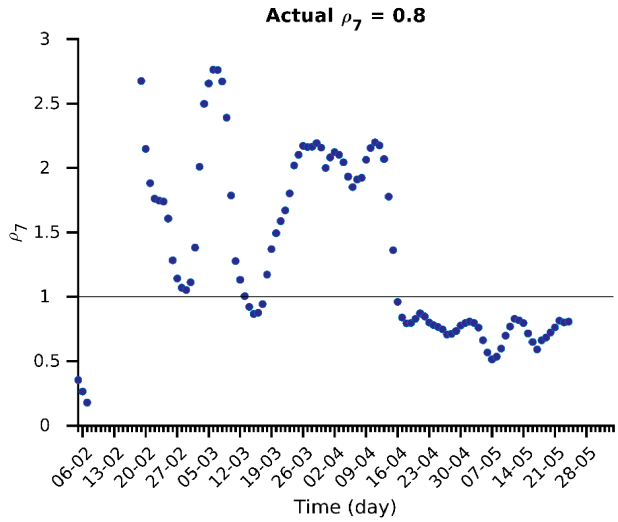


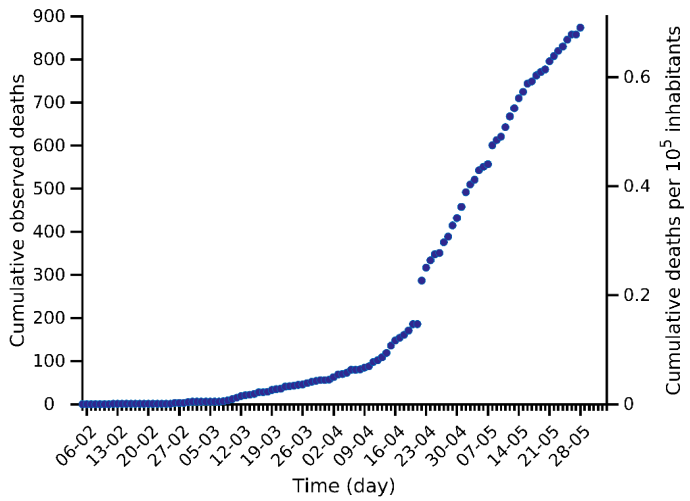


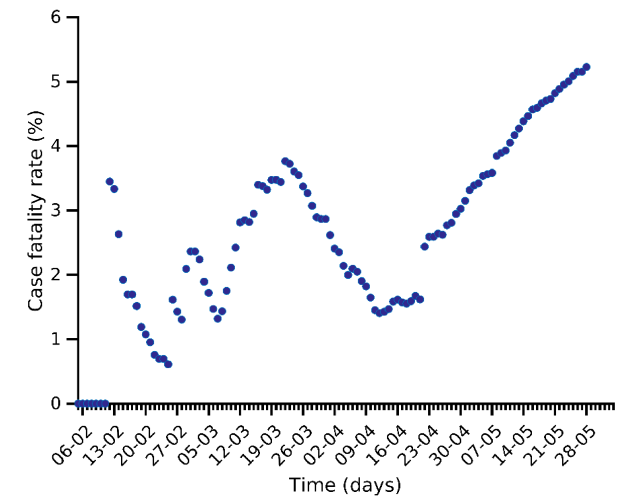




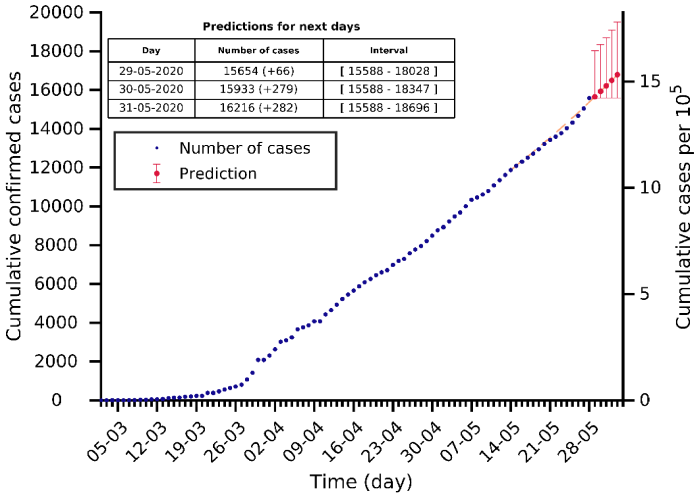


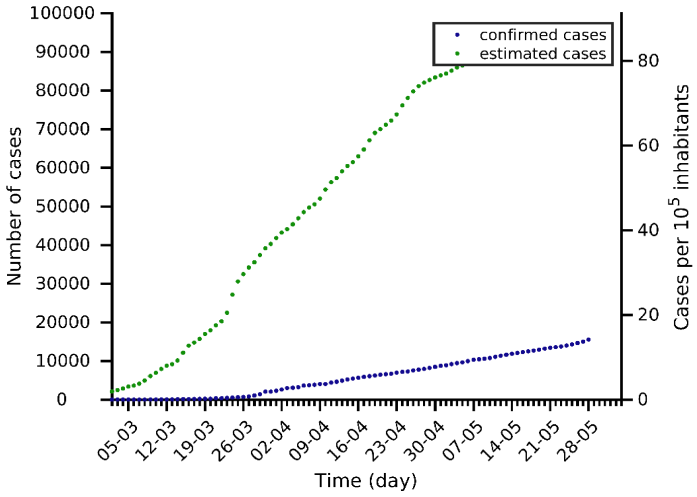


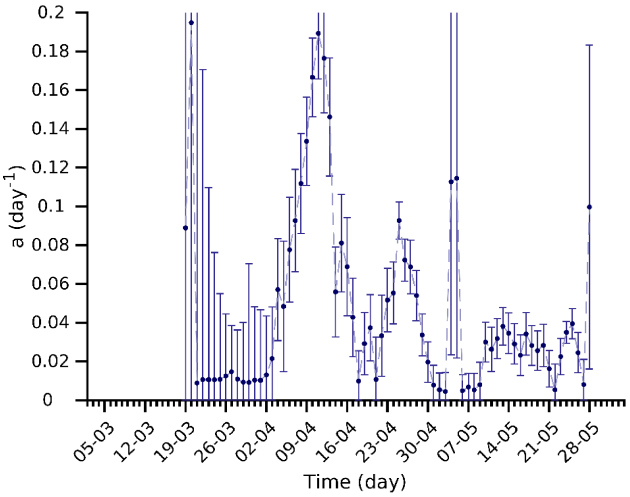


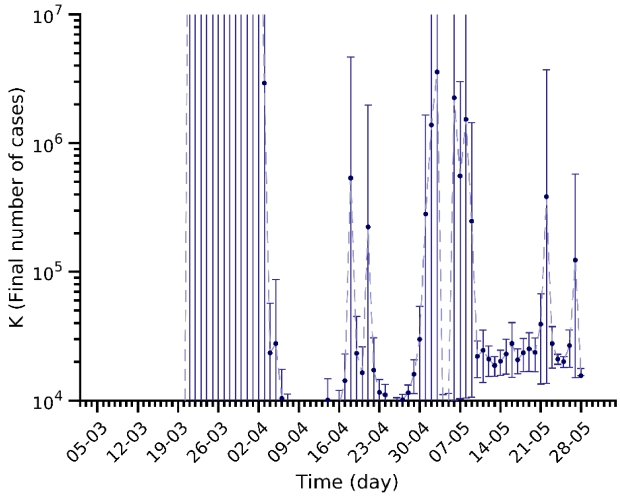


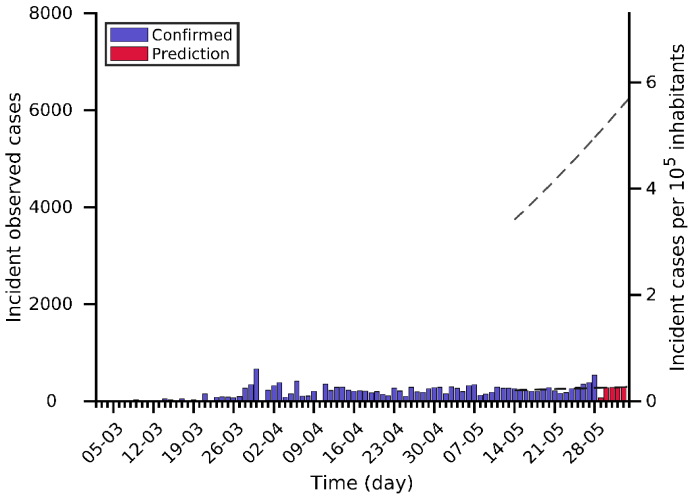
Philippines 28-05-2020. Population: 109.6M. Current cumulated incidence: $14/10^5$

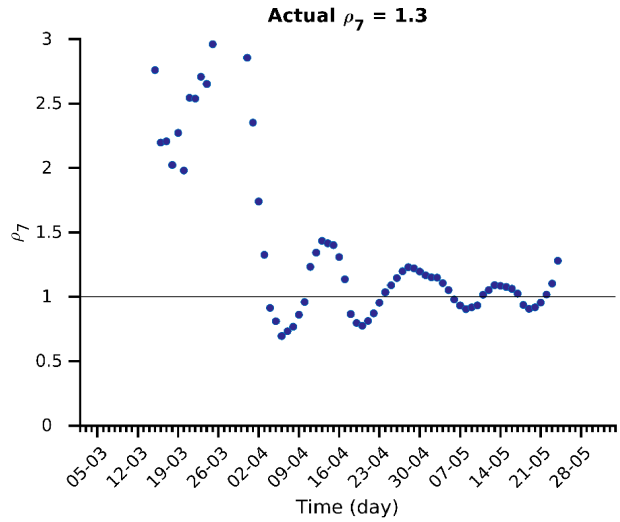


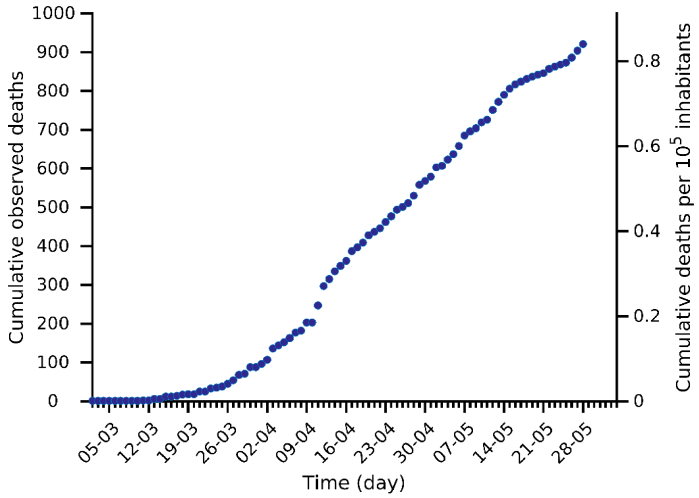


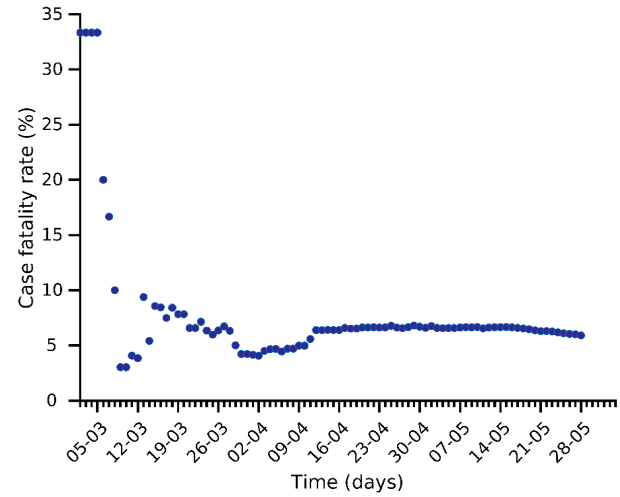




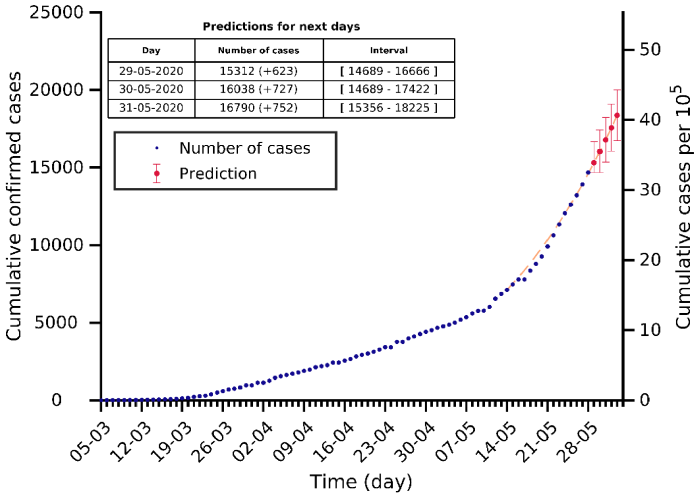


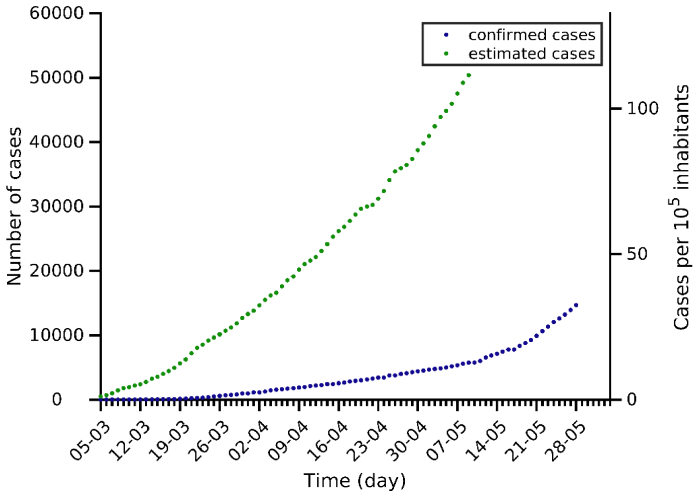


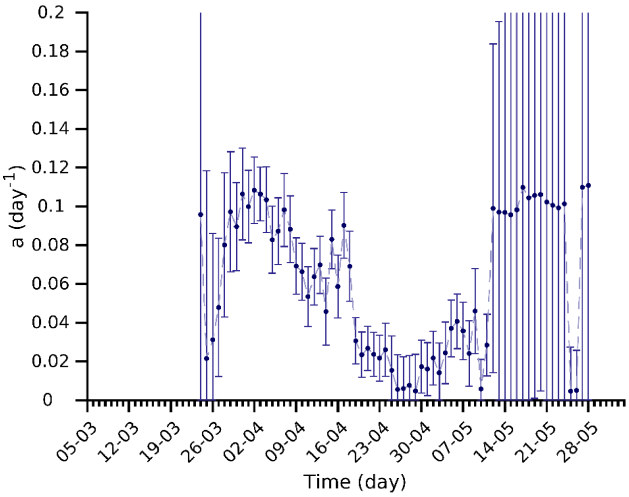


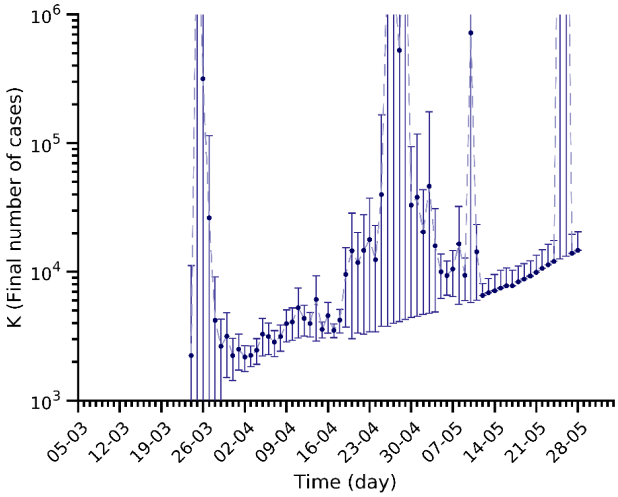


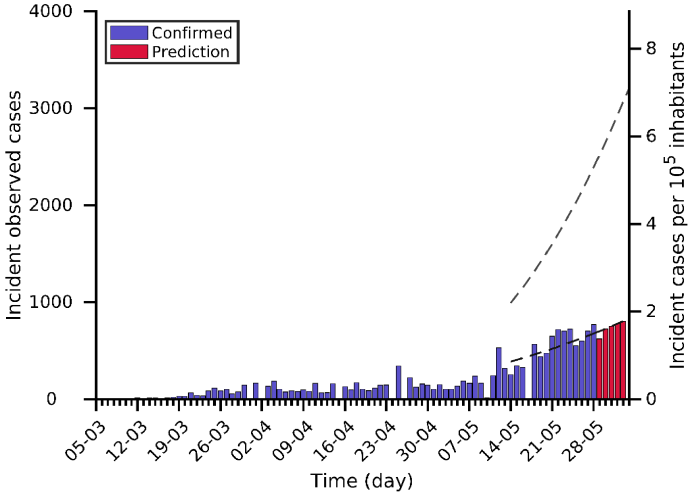
Argentina 28-05-2020. Population: 45.2M. Current cumulated incidence: 33/10^5

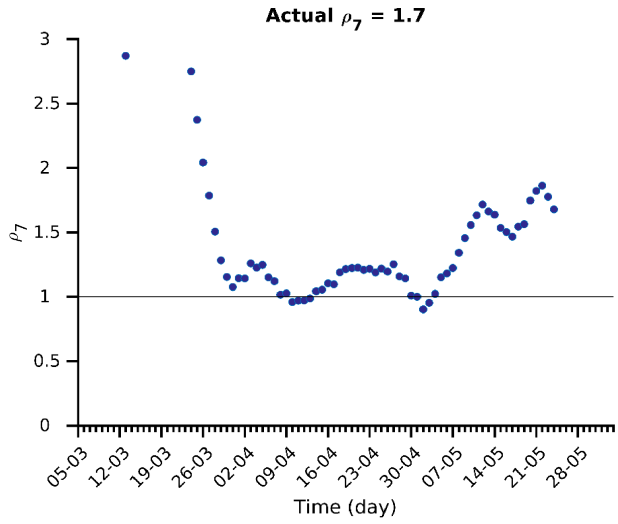


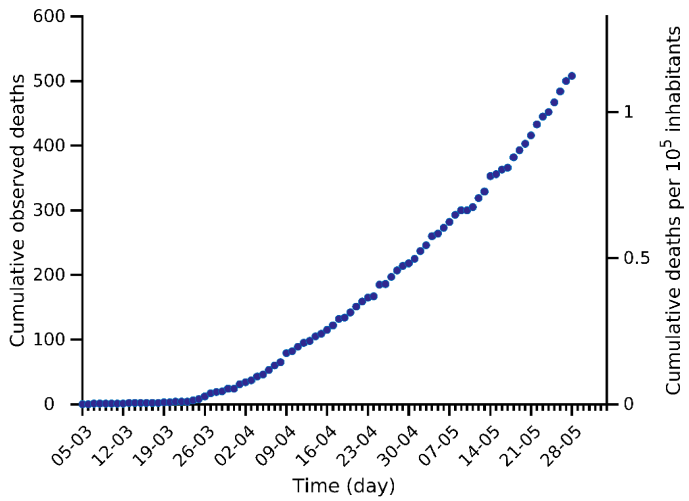


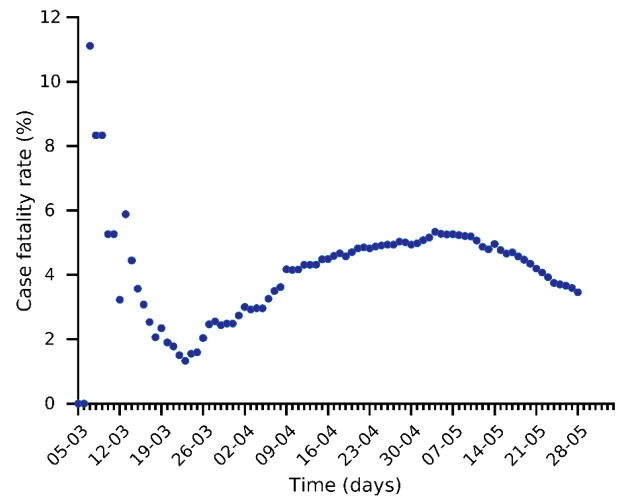




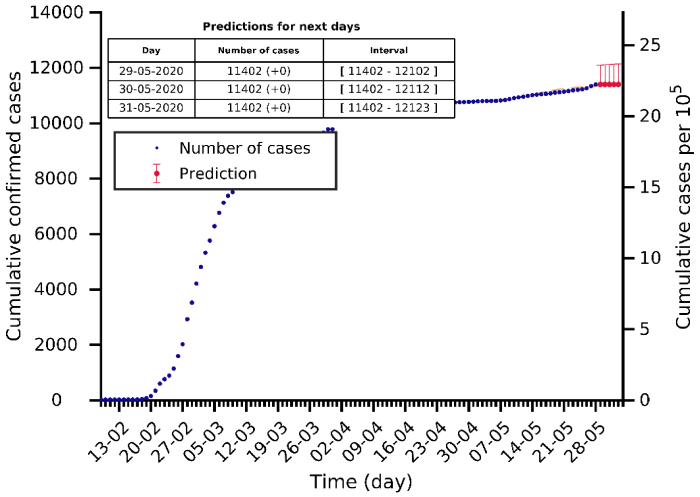


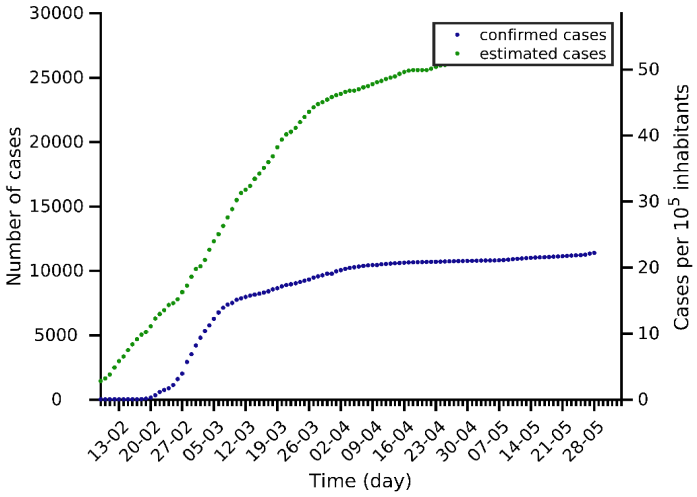


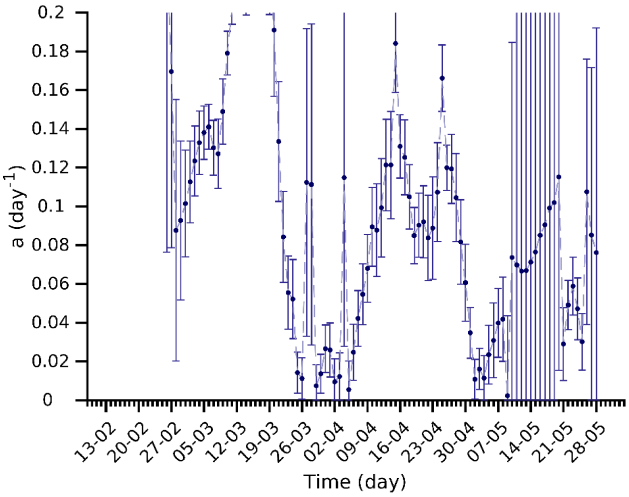


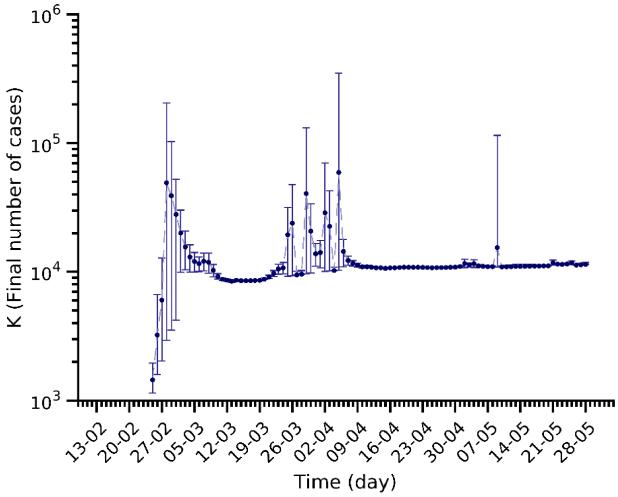


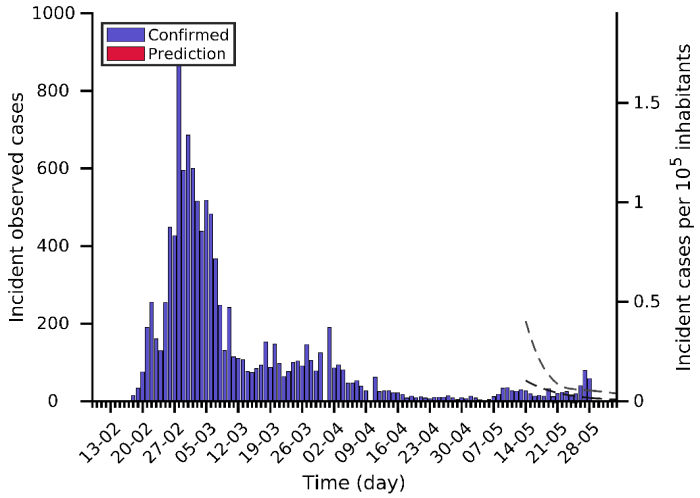
South Korea 28-05-2020. Population: 51.3M. Current cumulated incidence: $22/10^5$

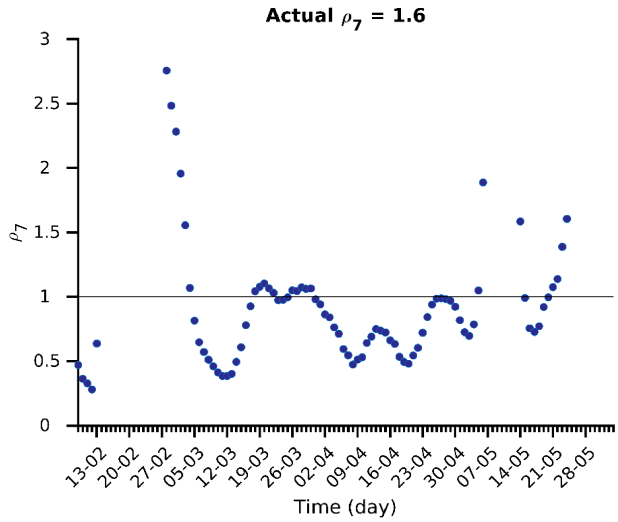


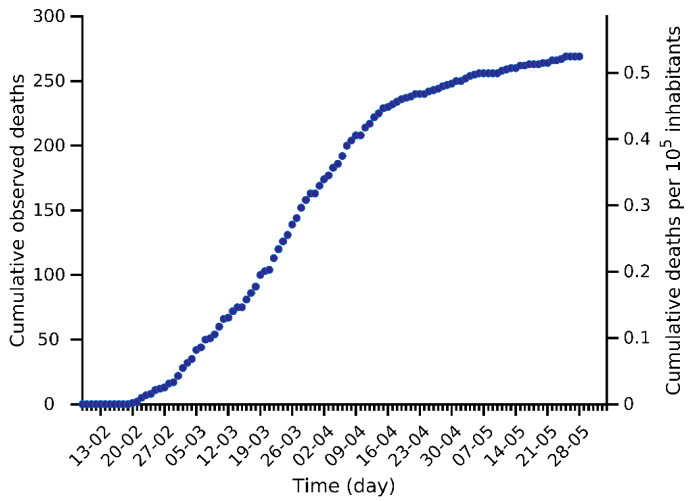


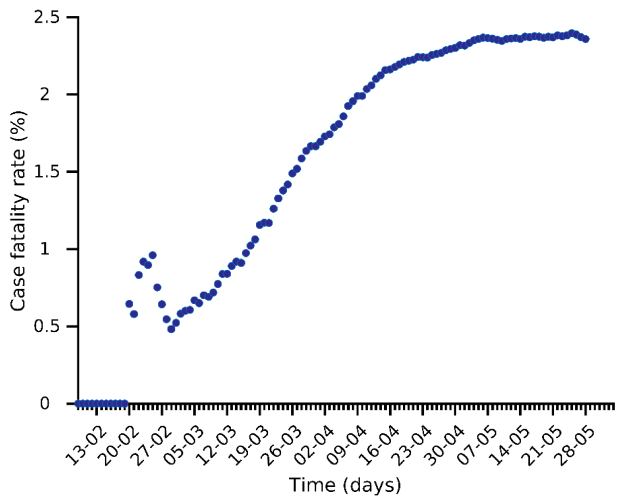




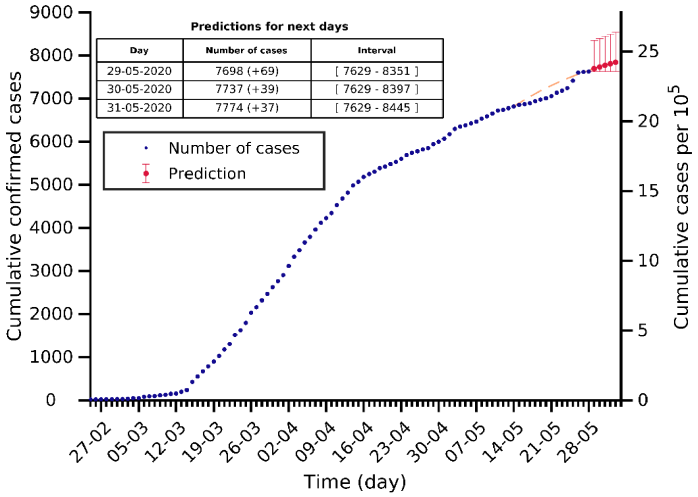


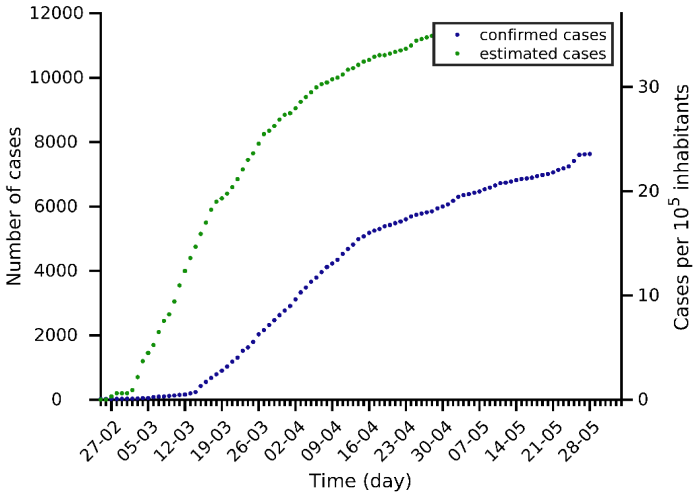


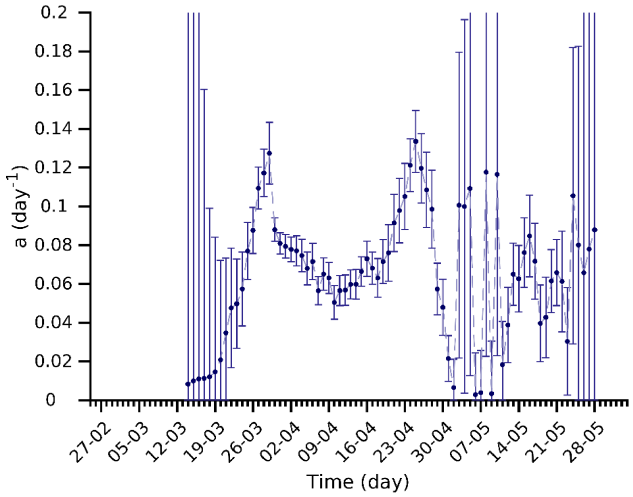


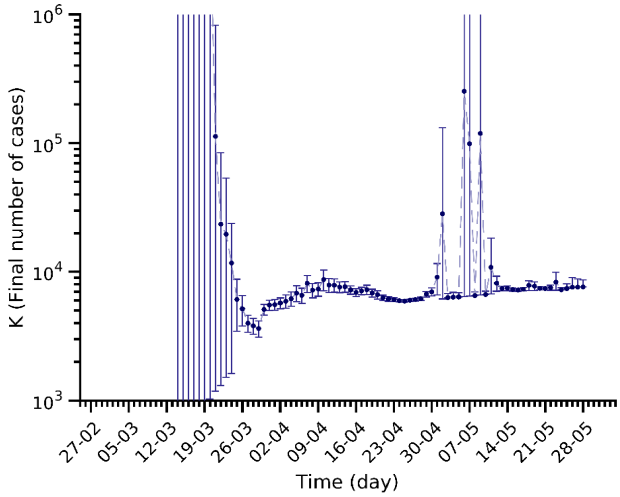


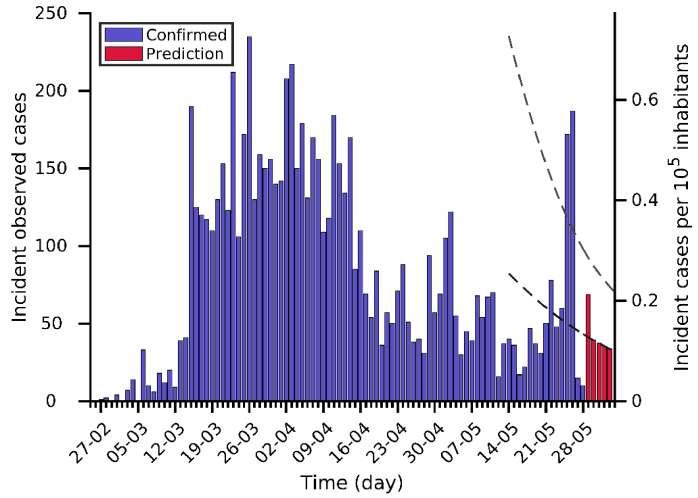
Malaysia 28-05-2020. Population: 32.4M. Current cumulated incidence: 24/10^5

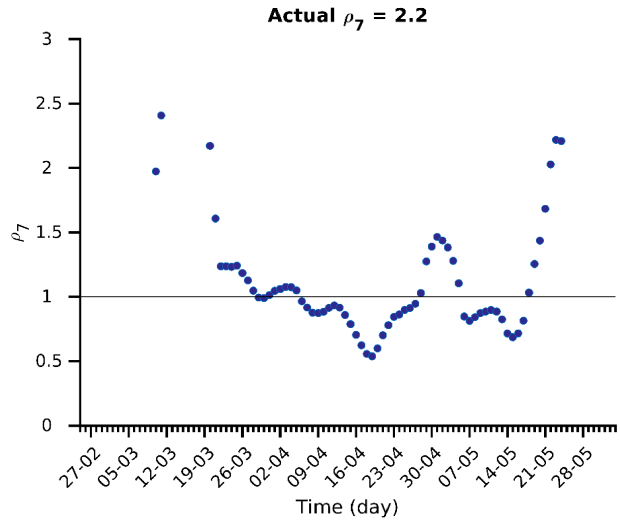


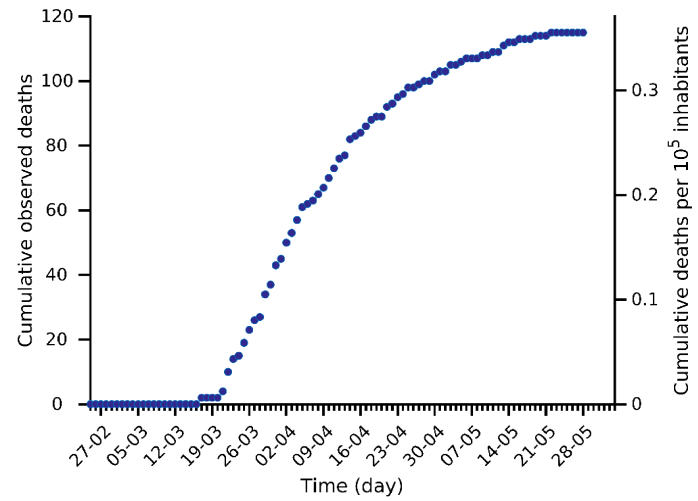


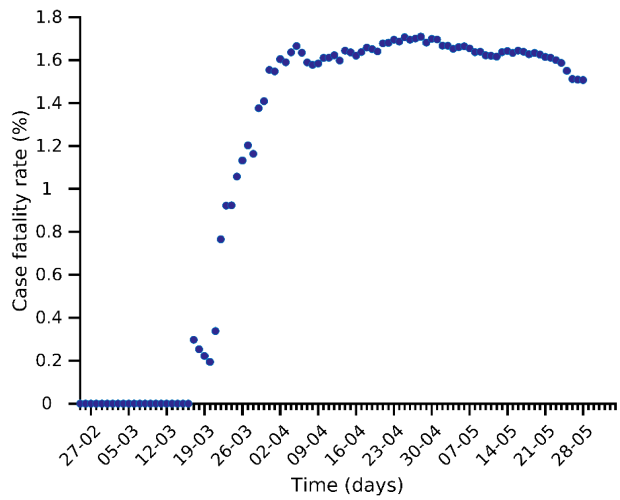




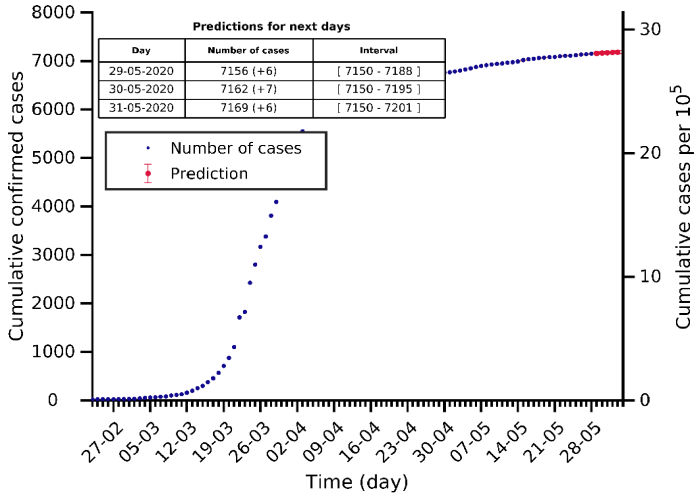


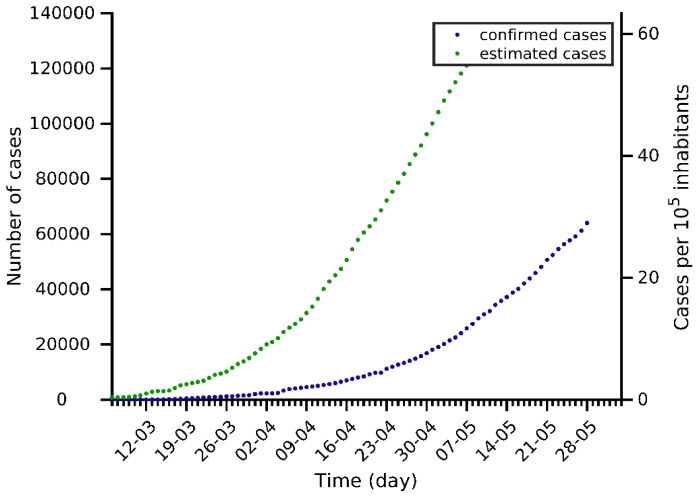


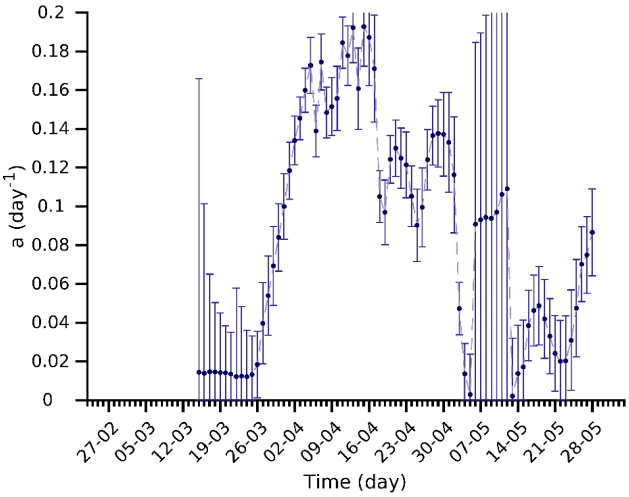


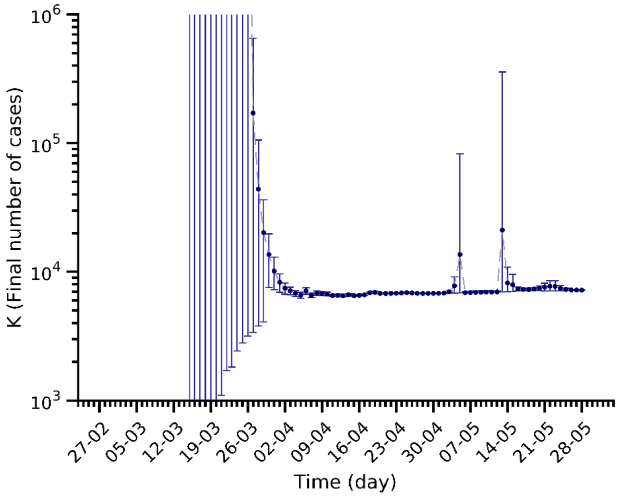


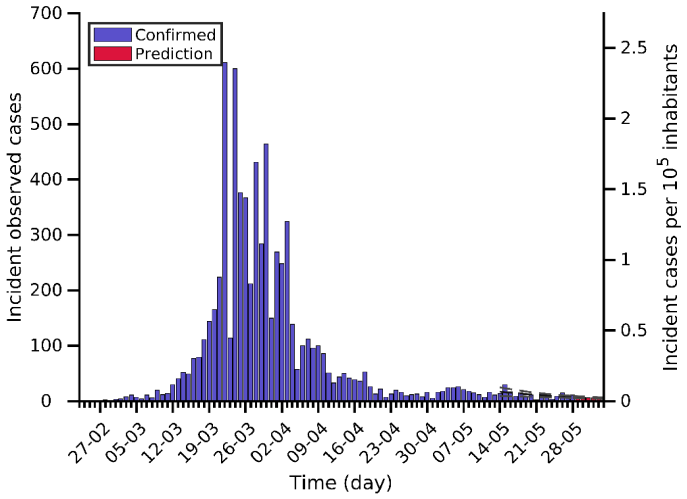
Australia 28-05-2020. Population: 25.5M. Current cumulated incidence: $28/10^5$

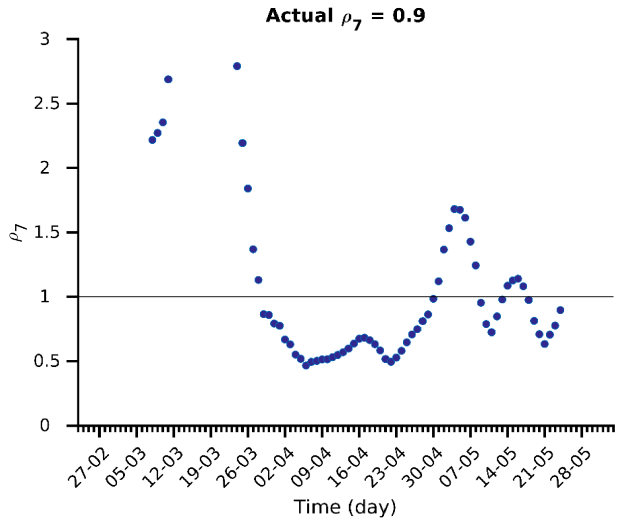


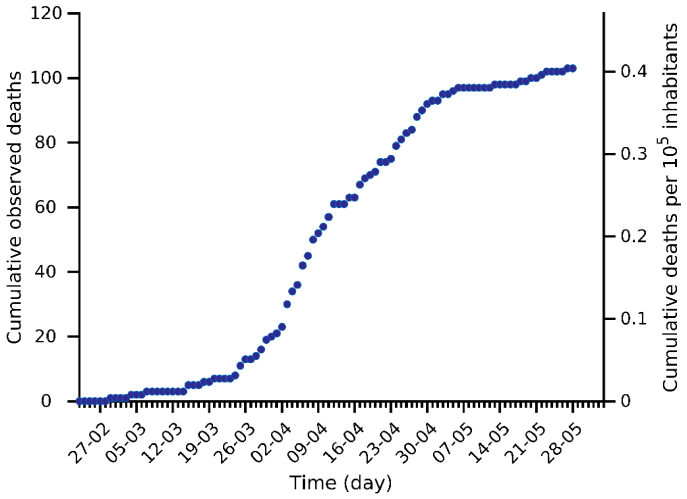


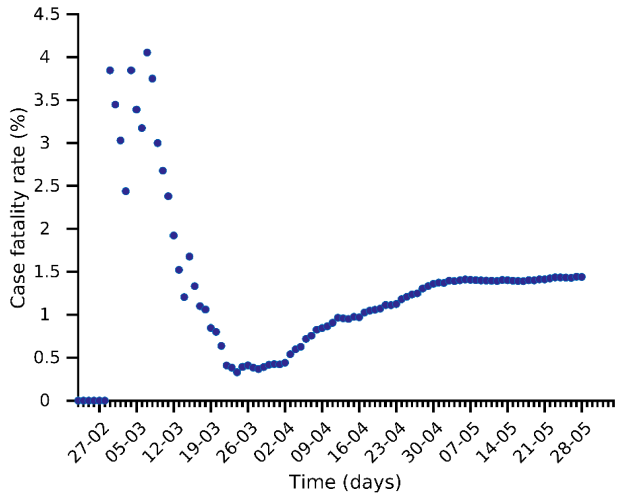




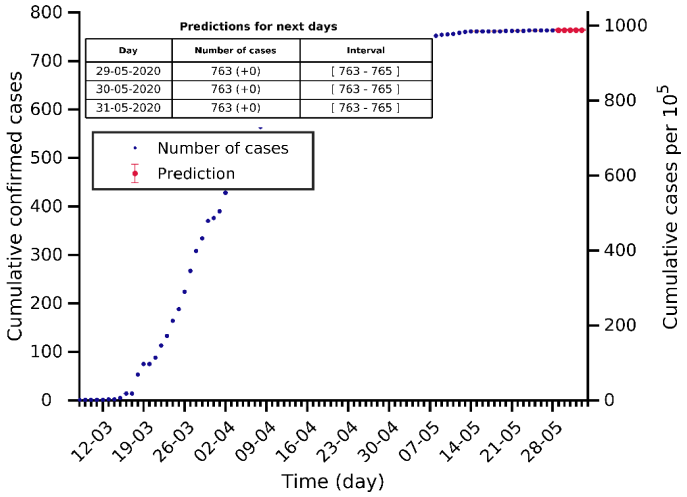


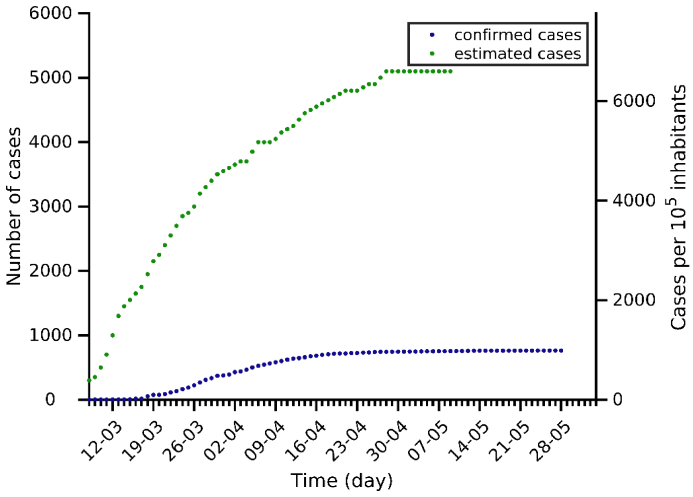


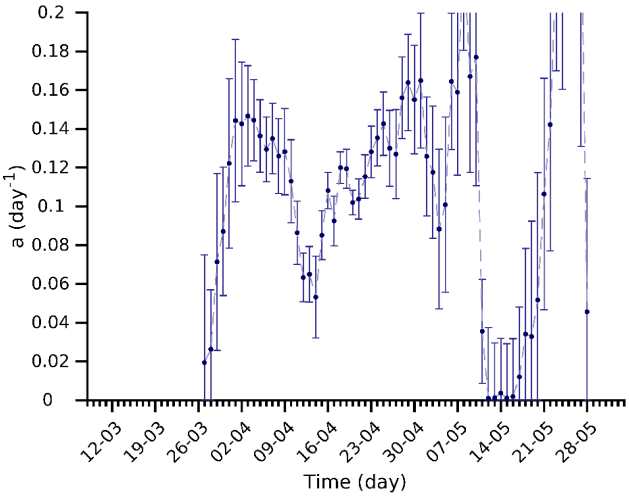


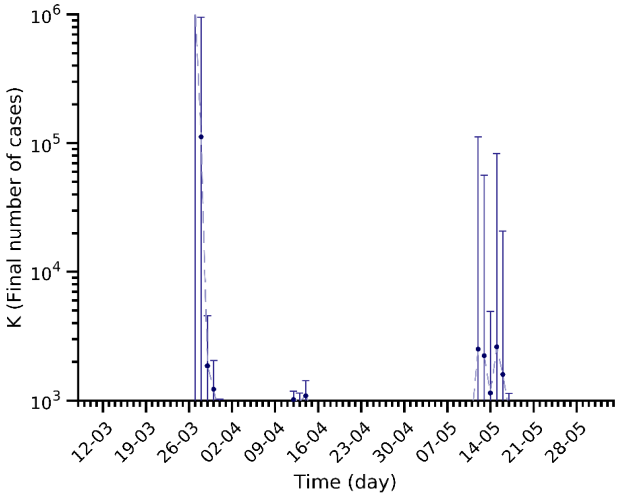


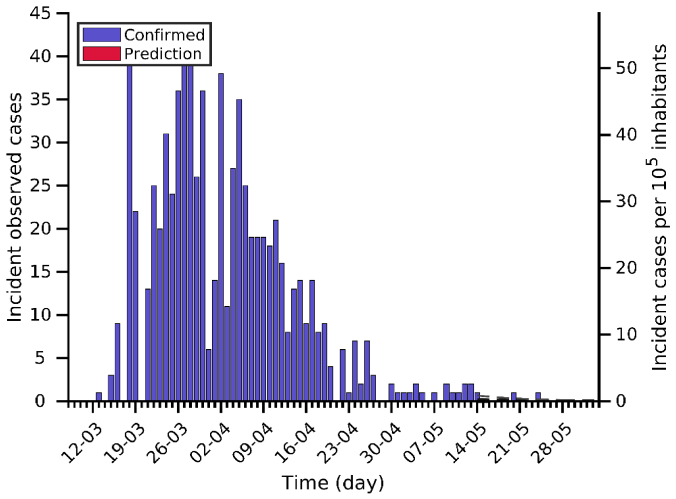
Andorra 28-05-2020. Population: 0.1M. Current cumulated incidence: 988/10^5

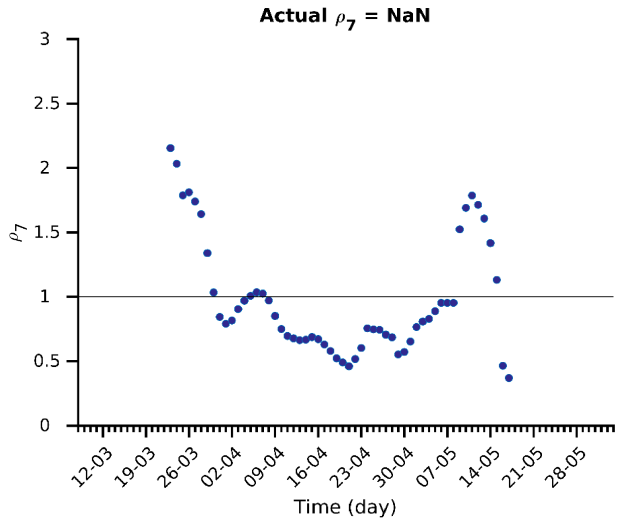


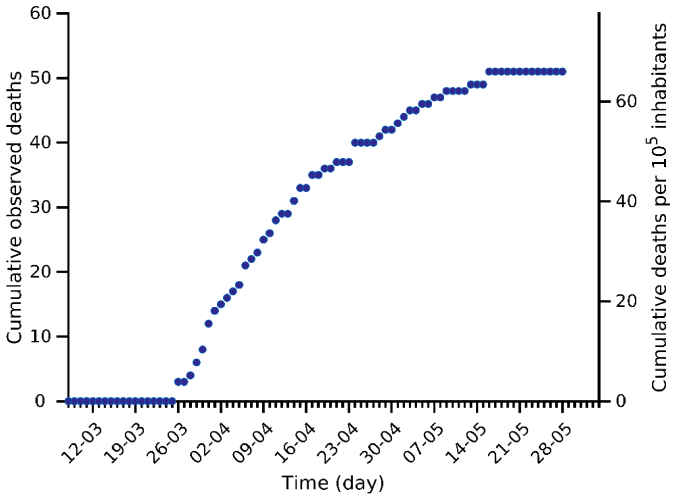


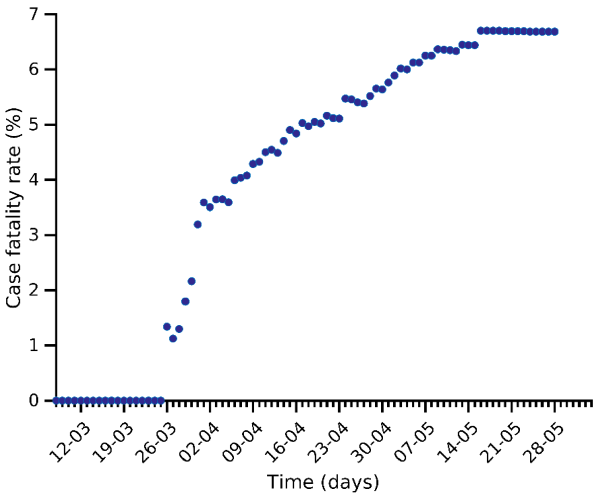








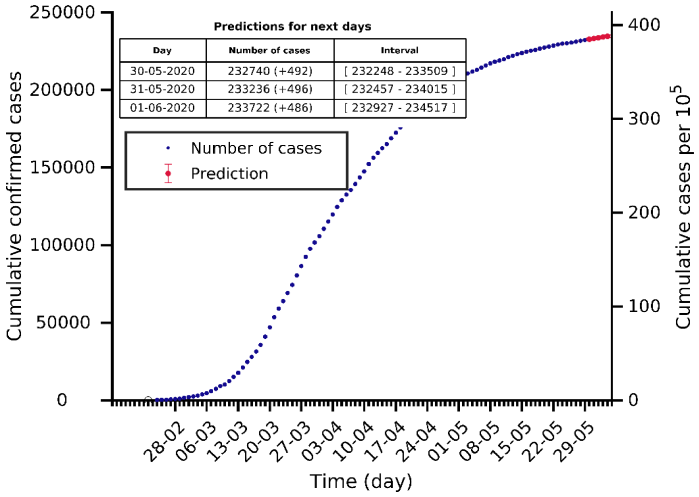


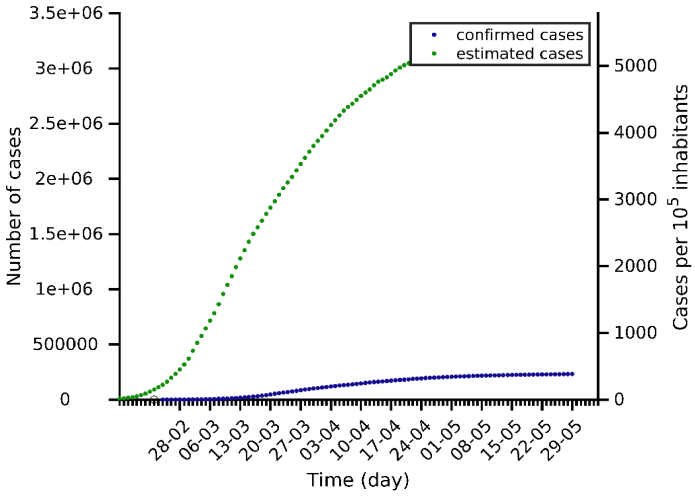


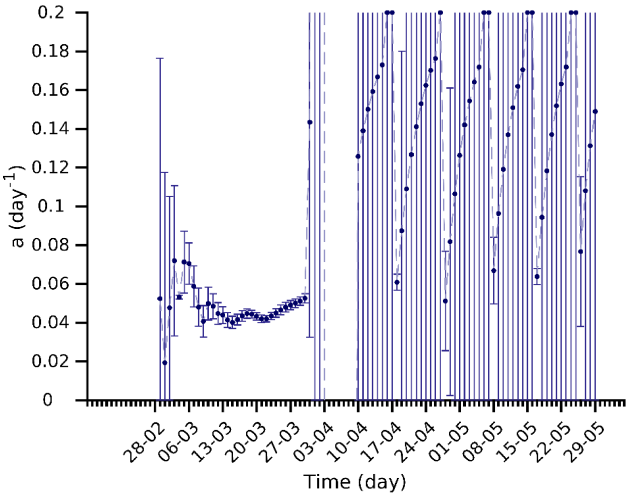
(3) Analysis and prediction of COVID-19 for Italy and its regions

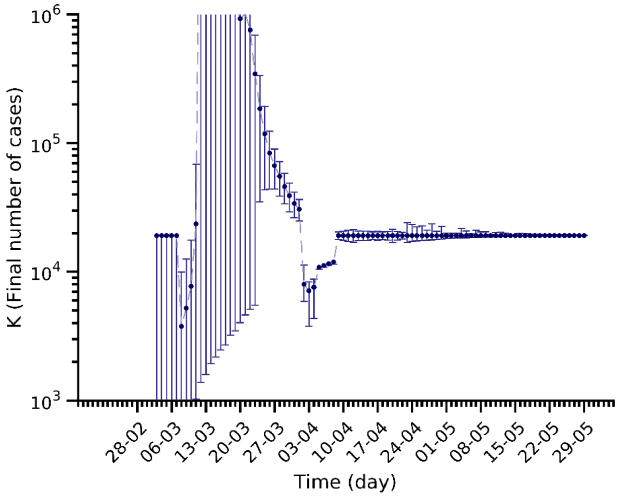
Data obtained from: https://github.com/pcm-dpc/COVID-19/tree/master/dati-andamento-nazionale

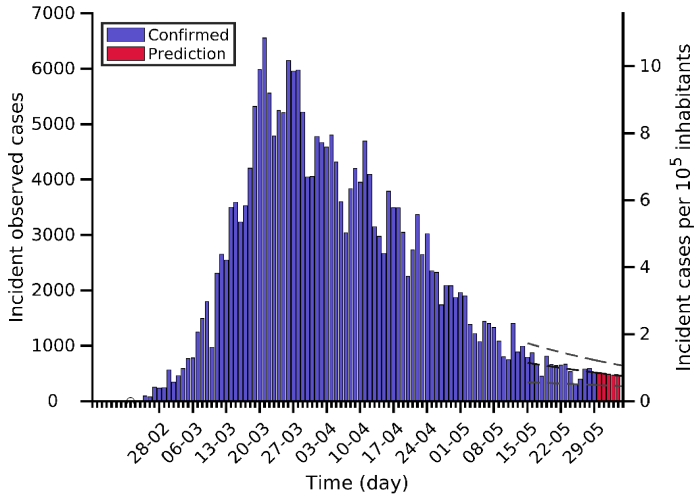
Italy 29-05-2020. Population: 60.5M. Current cumulated incidence: $384/10^5$

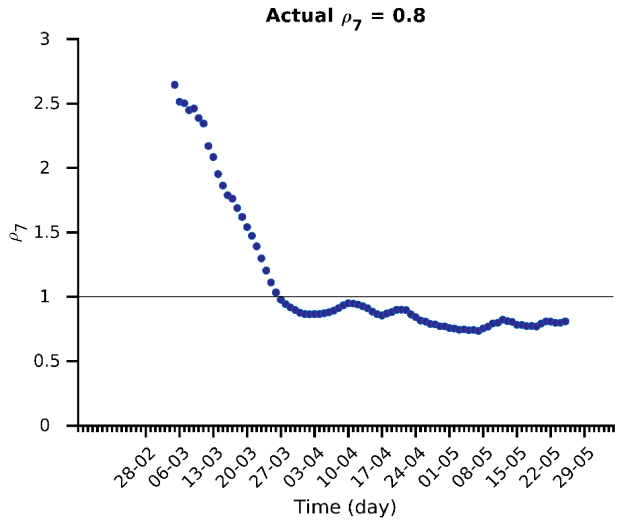


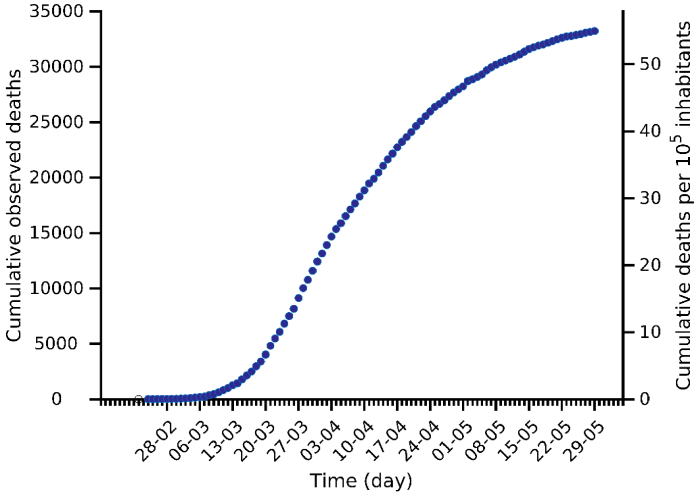


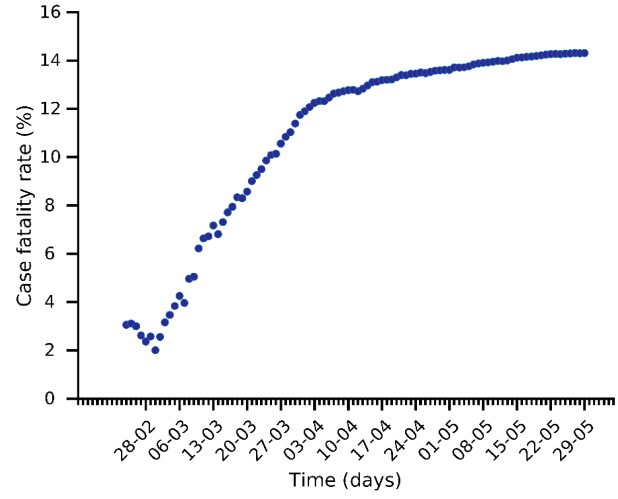




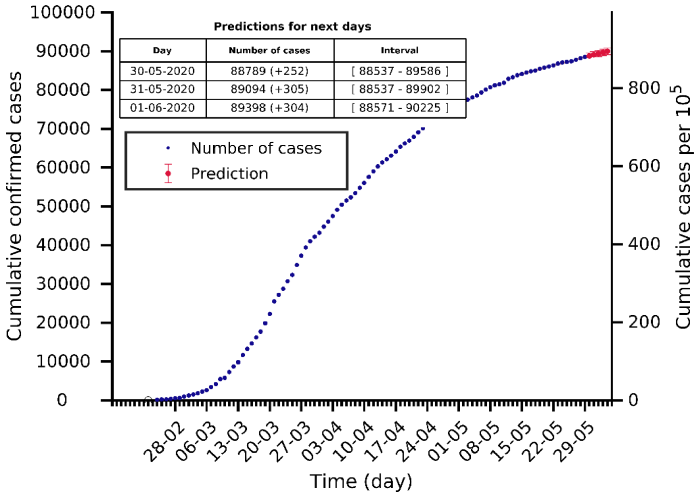


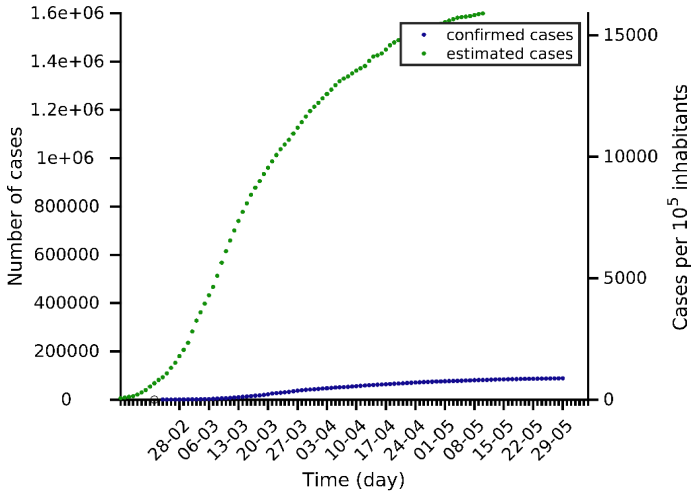


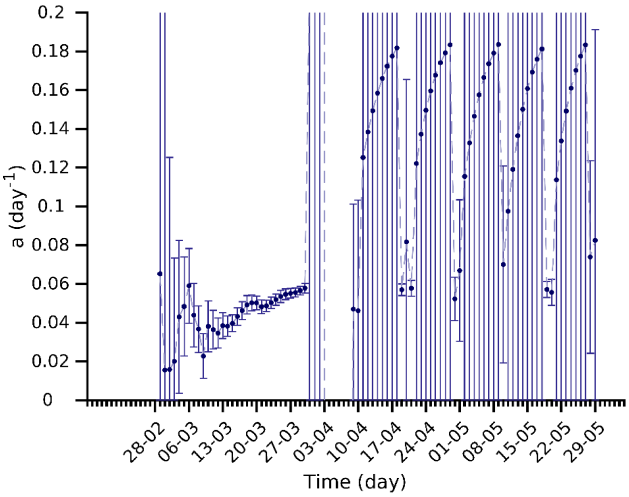


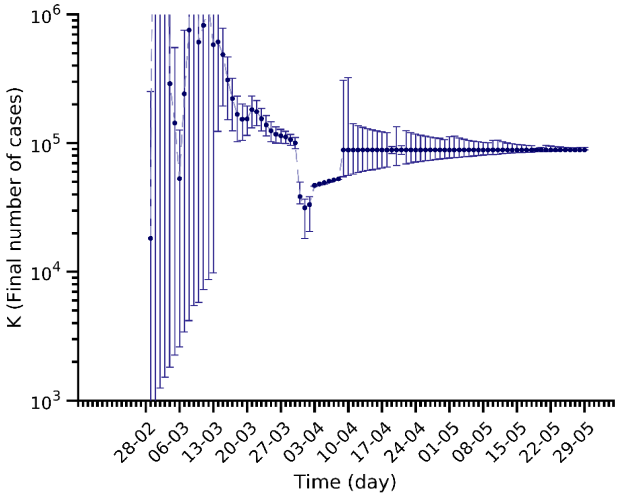


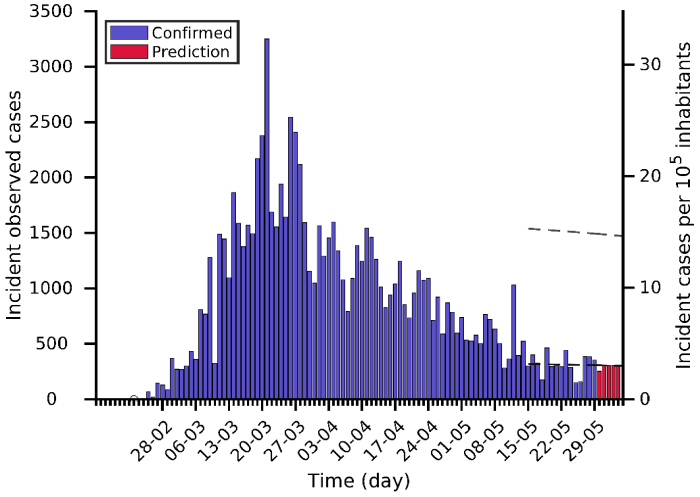
Lombardia 29-05-2020. Population: 10.1M. Current cumulated incidence: 880/10^5

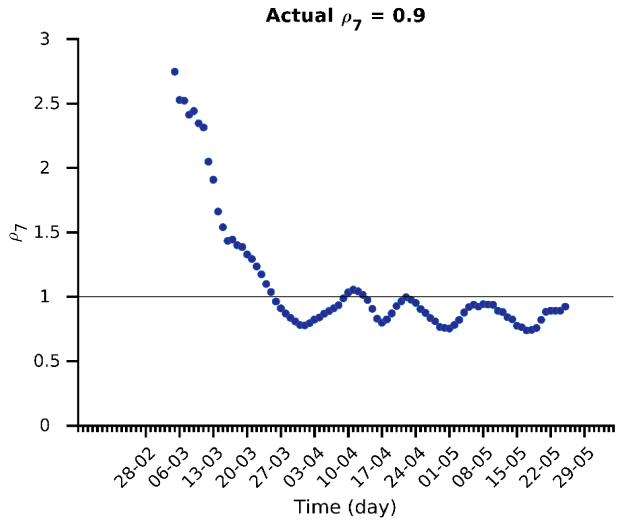


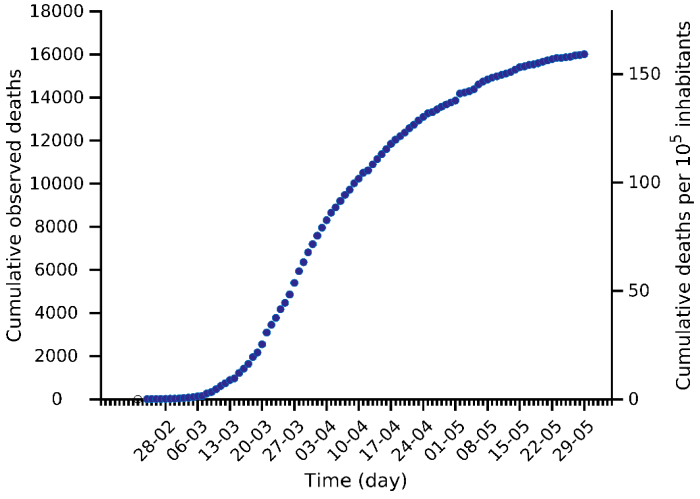


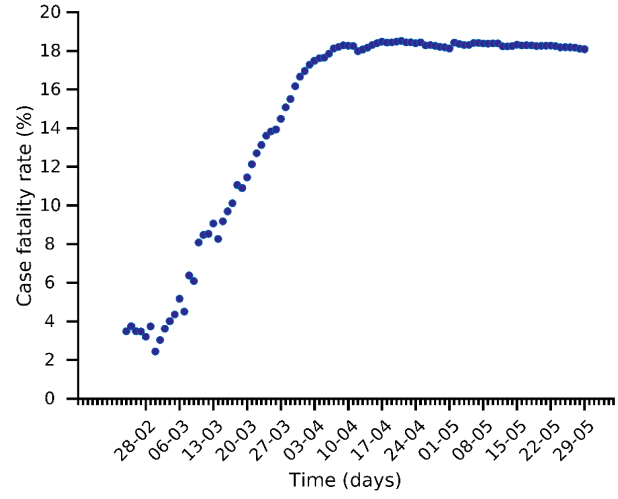




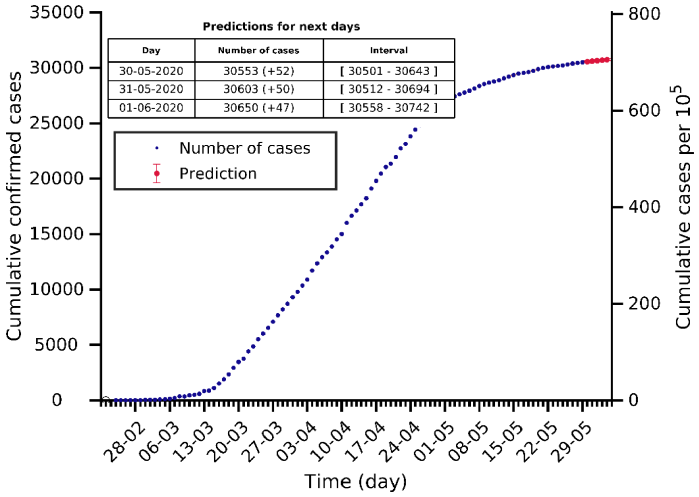


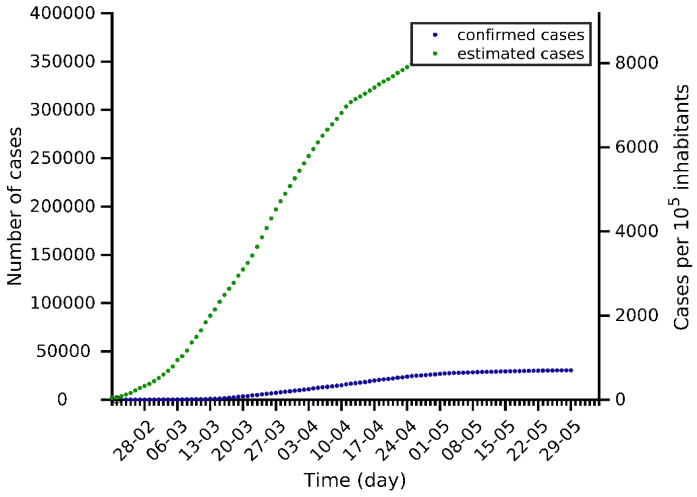


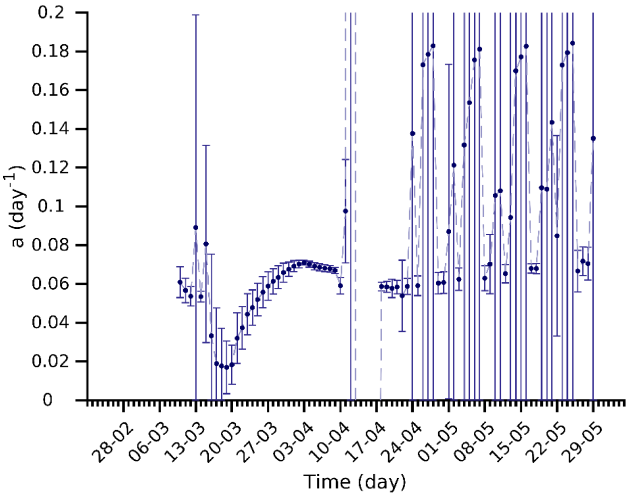


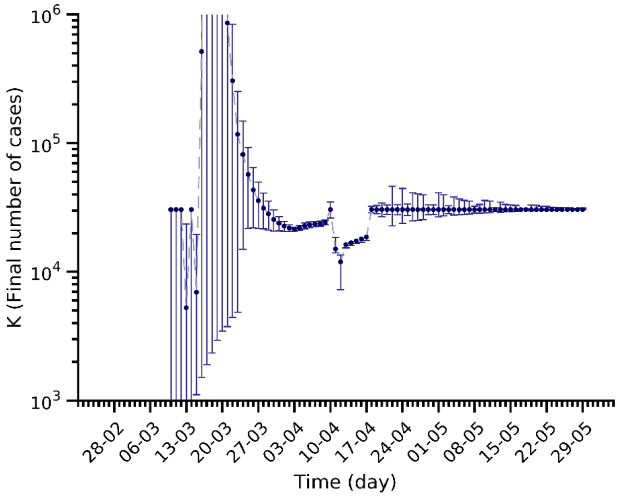


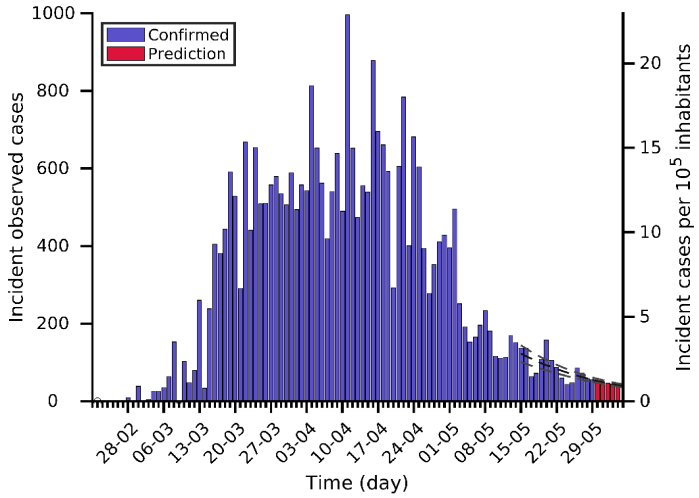
Piemonte 29-05-2020. Population: 4.4M. Current cumulated incidence: 700/10⁵

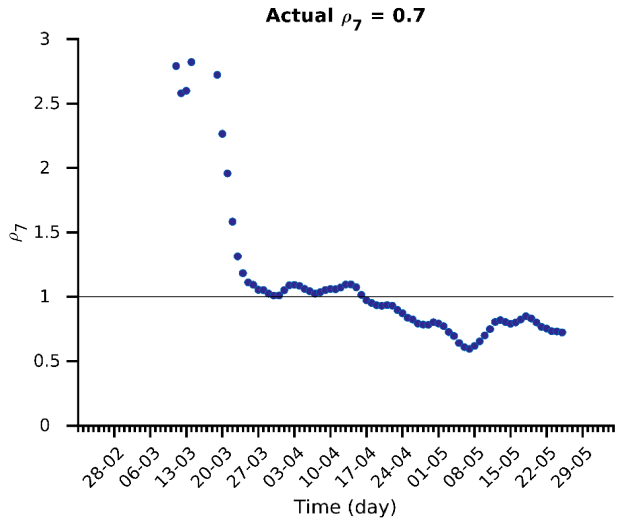


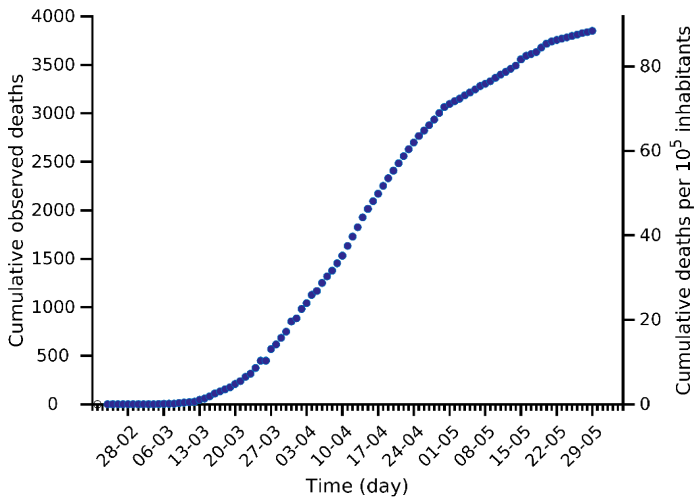


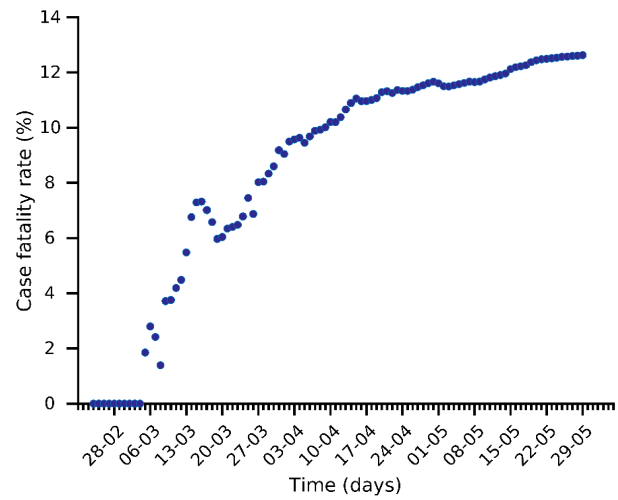




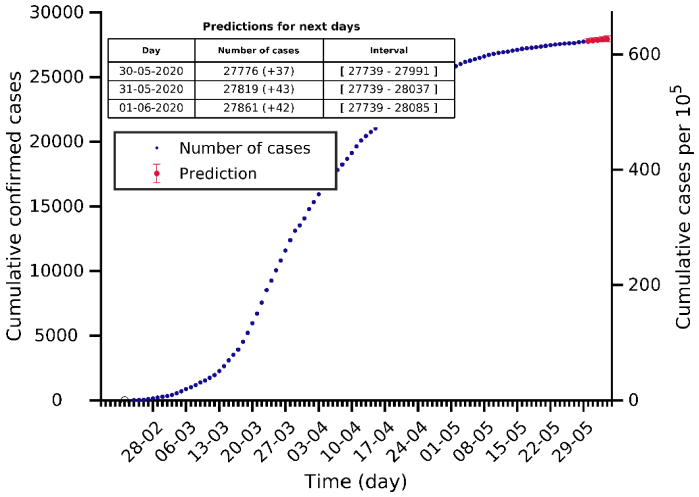


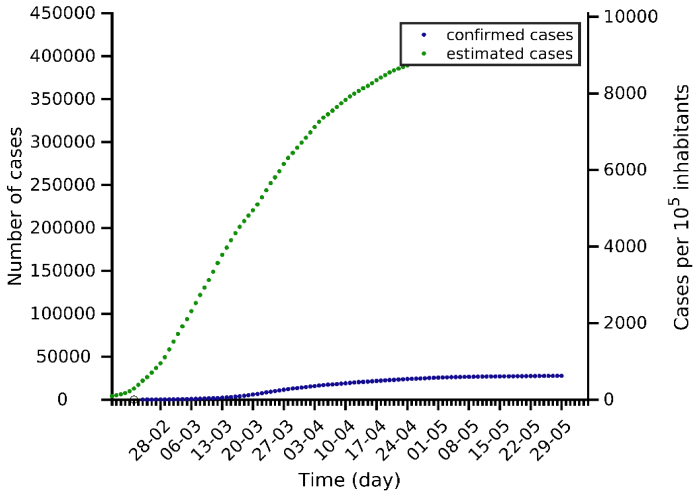


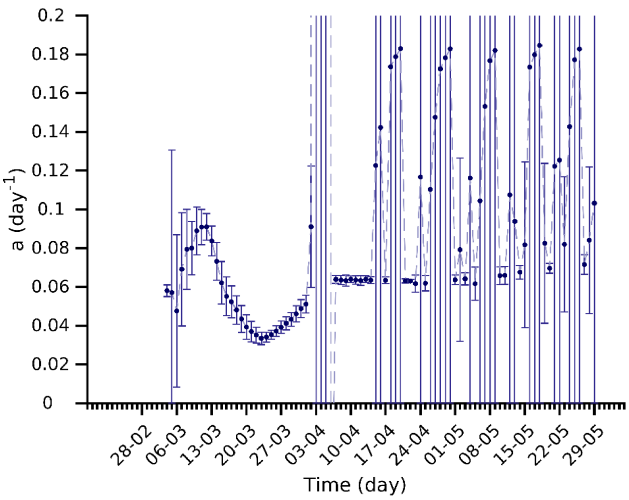


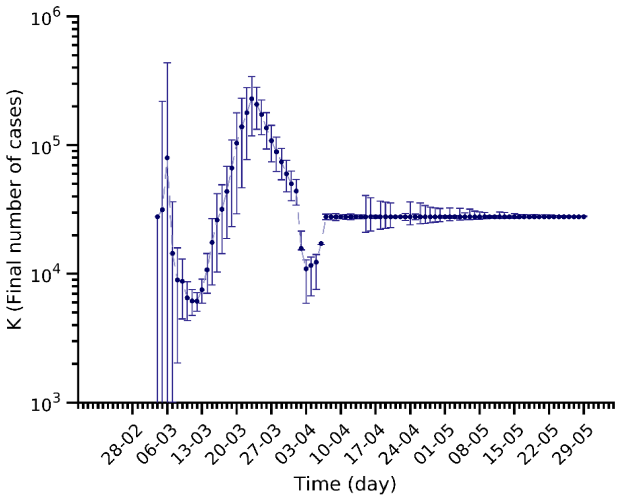


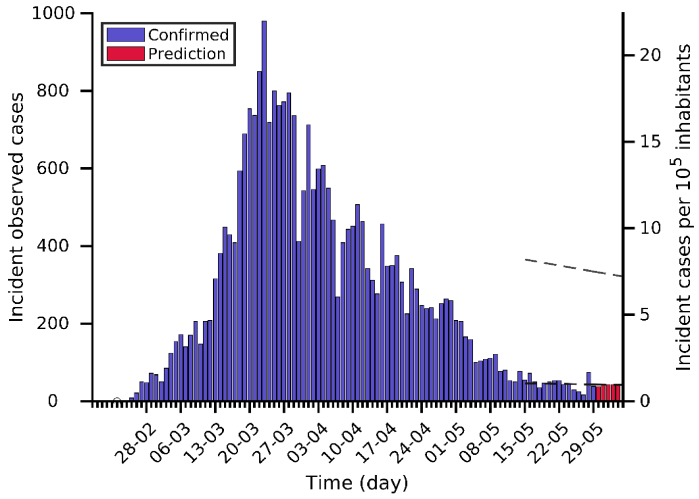
Emilia Romagna 29-05-2020. Population: 4.5M. Current cumulated incidence: 622/1

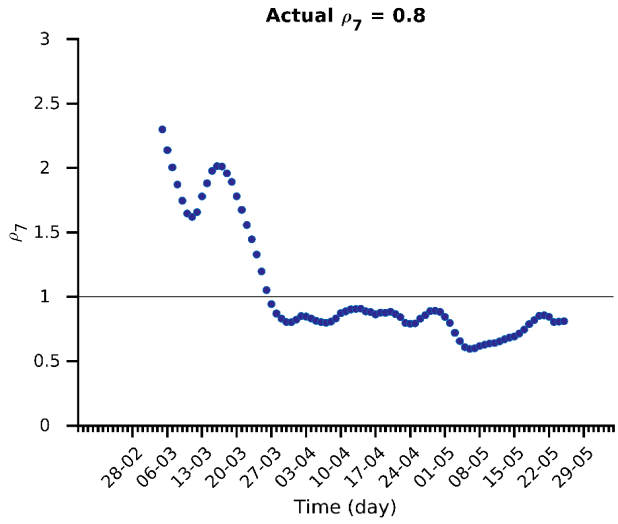


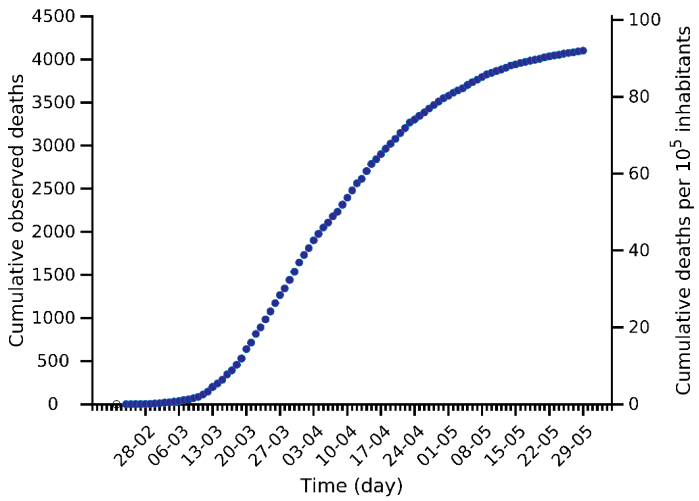


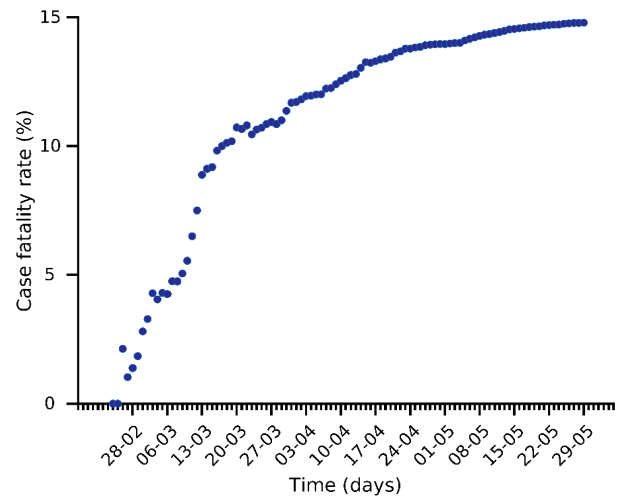




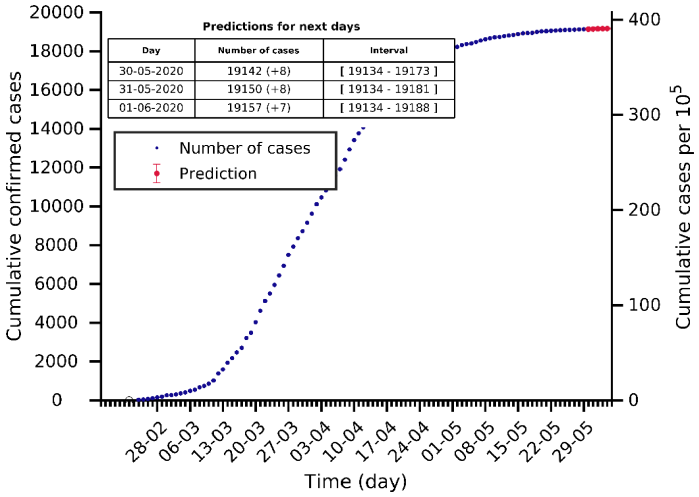


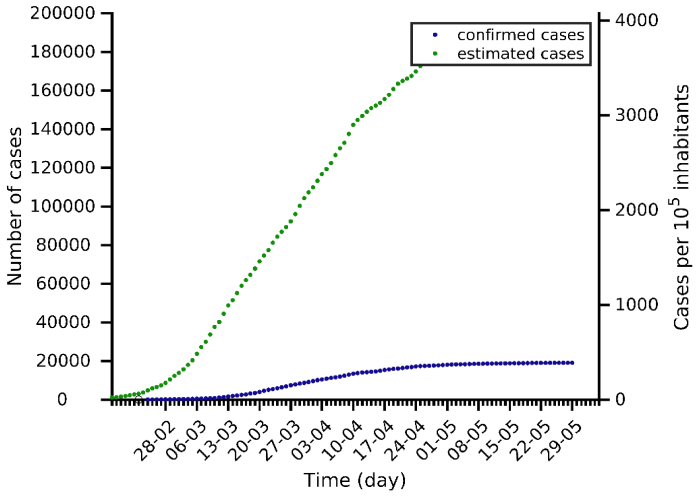


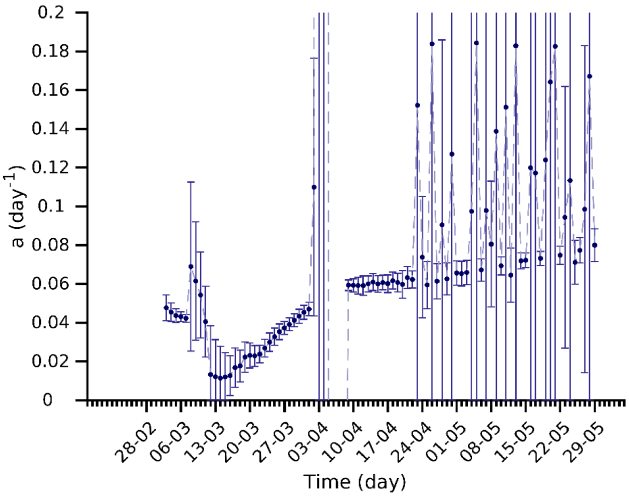


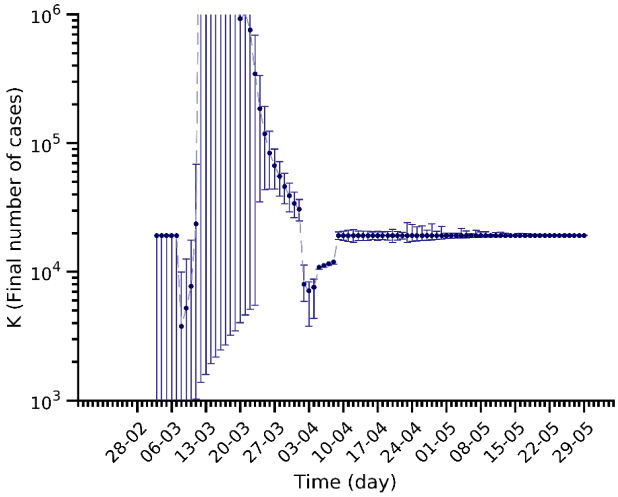


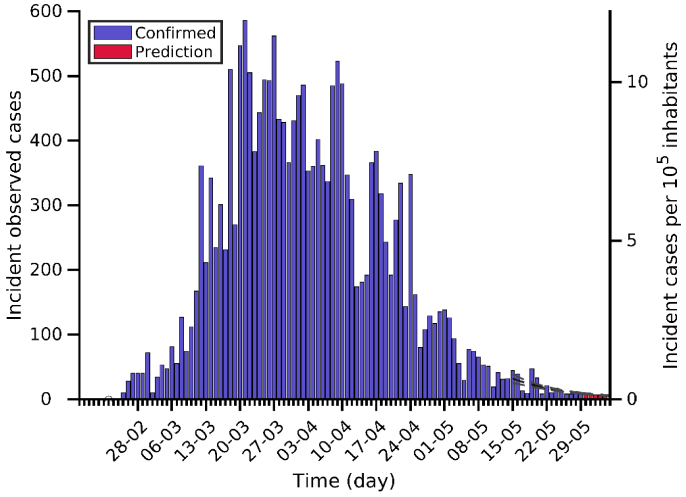
Veneto 29-05-2020. Population: 4.9M. Current cumulated incidence: 390/10⁵

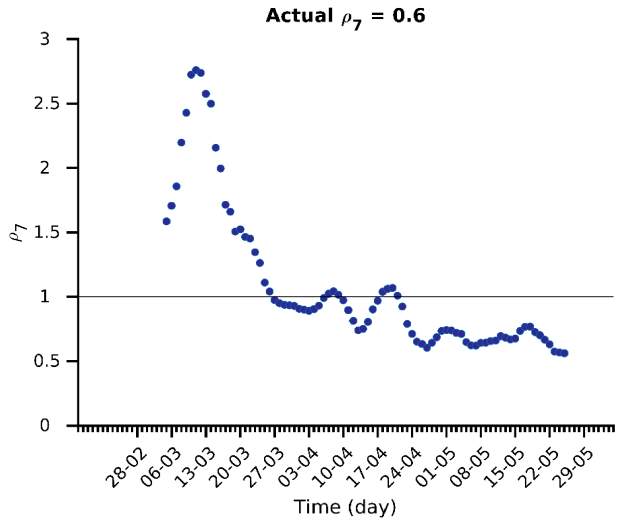


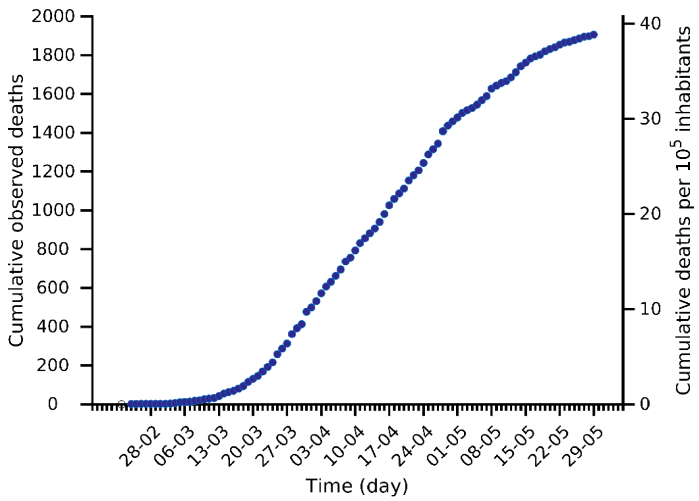


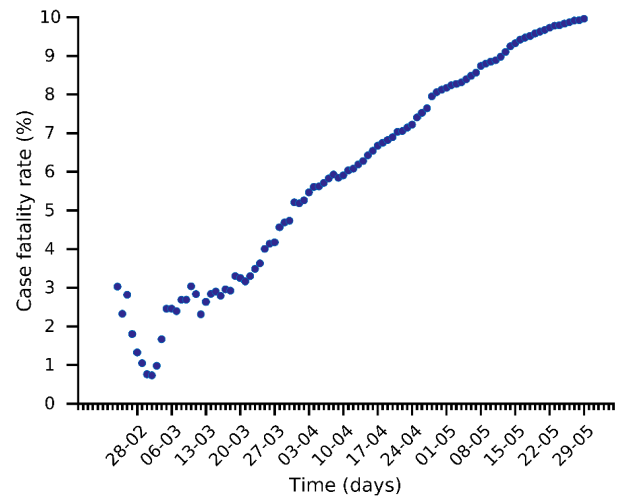




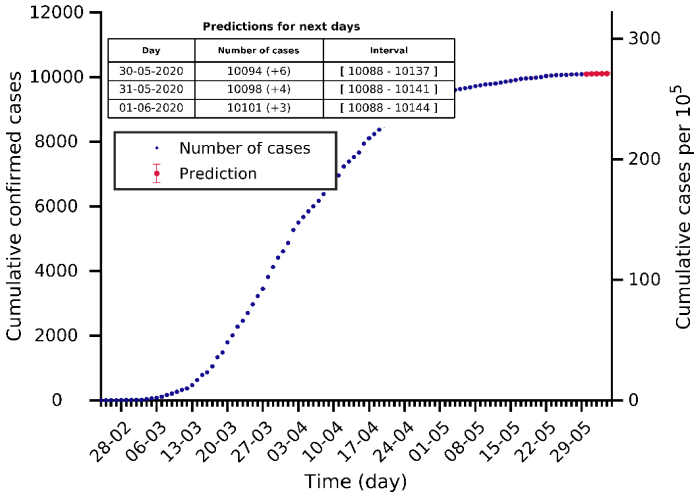


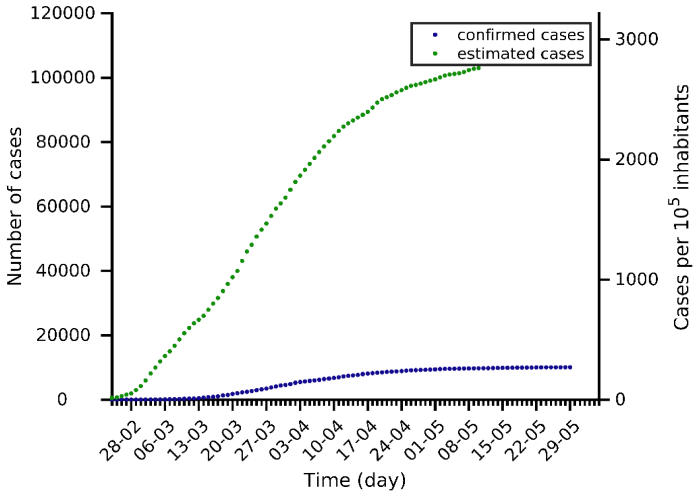


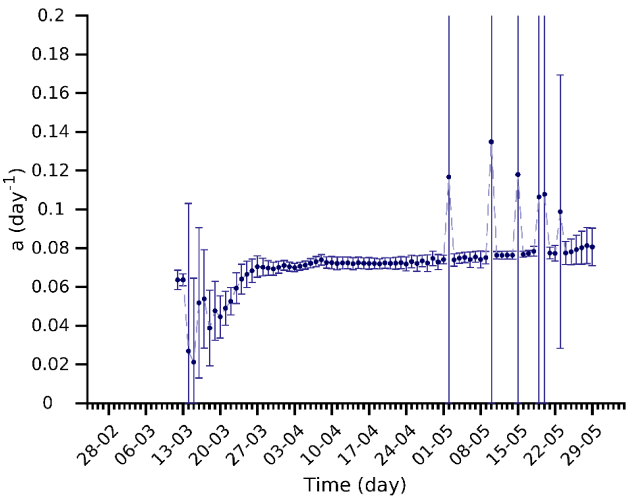


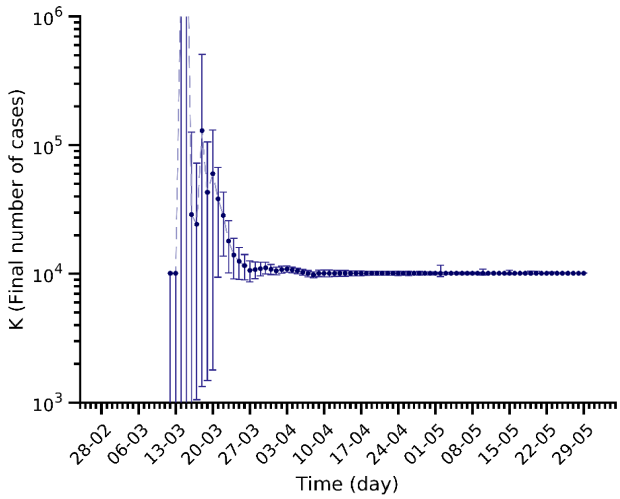


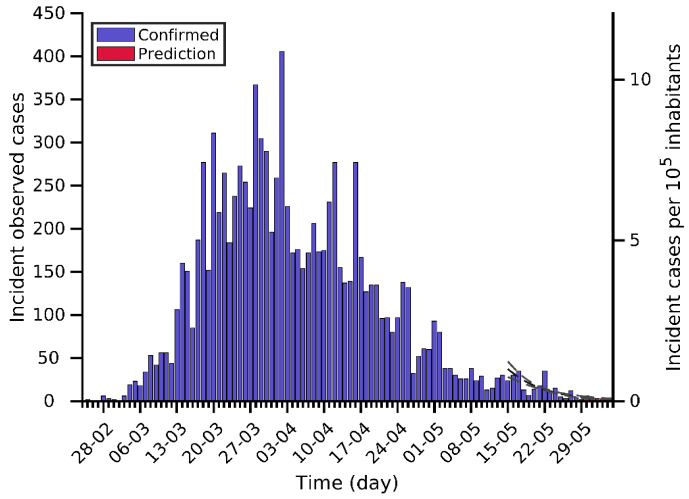
Toscana 29-05-2020. Population: 3.7M. Current cumulated incidence: 270/10⁵

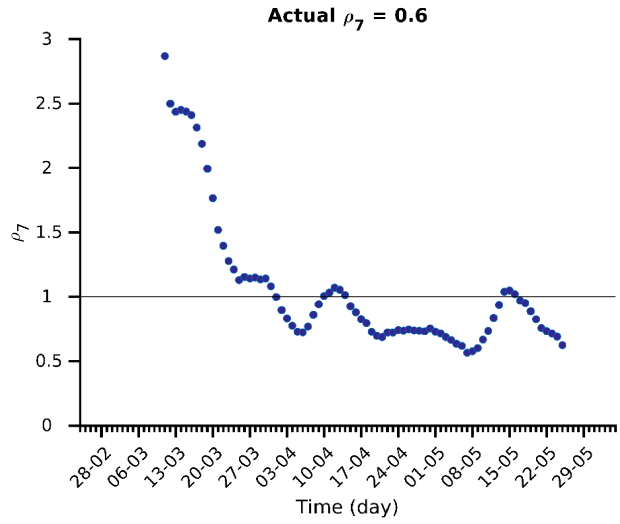


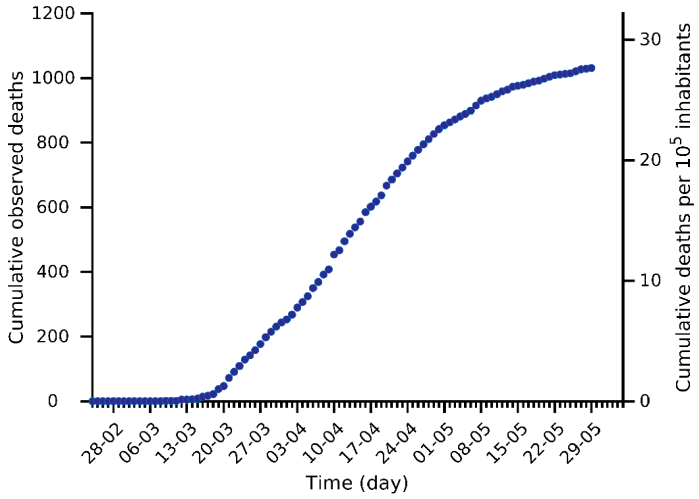


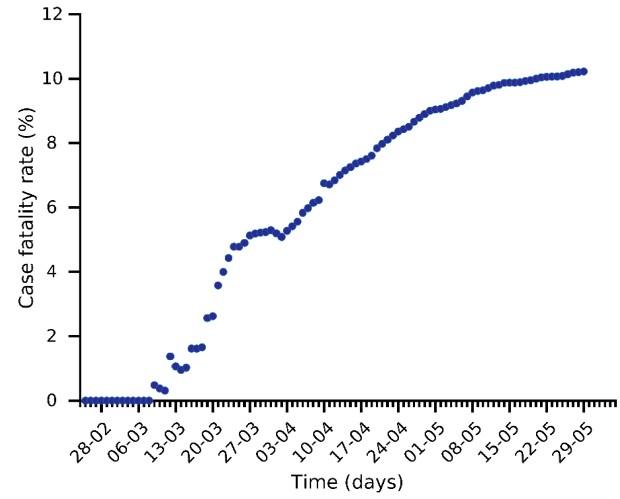




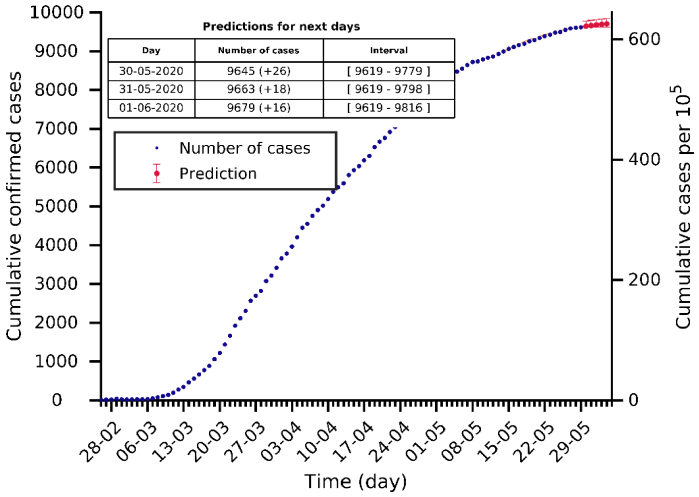


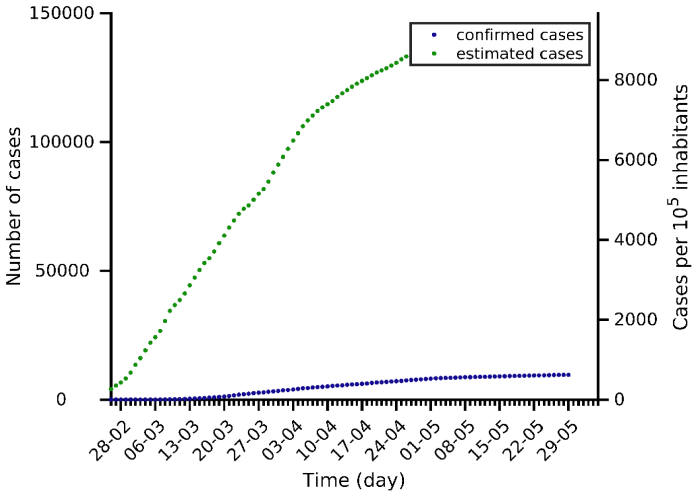


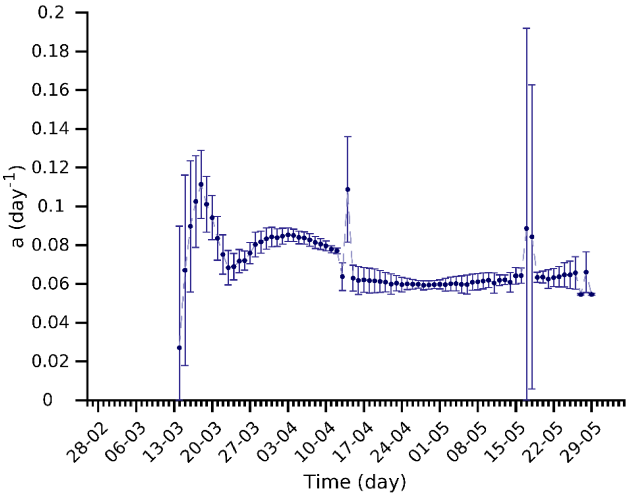


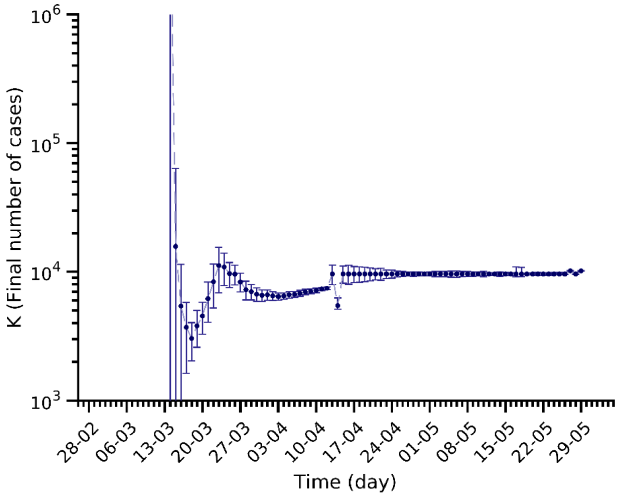


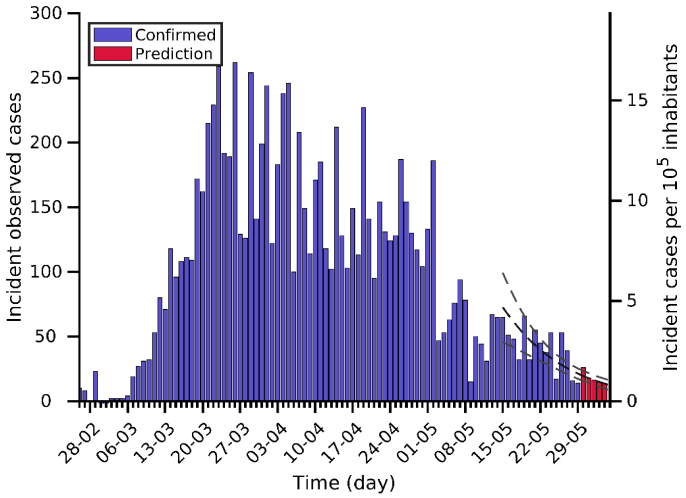
Liguria 29-05-2020. Population: 1.6M. Current cumulated incidence: 620/10⁵

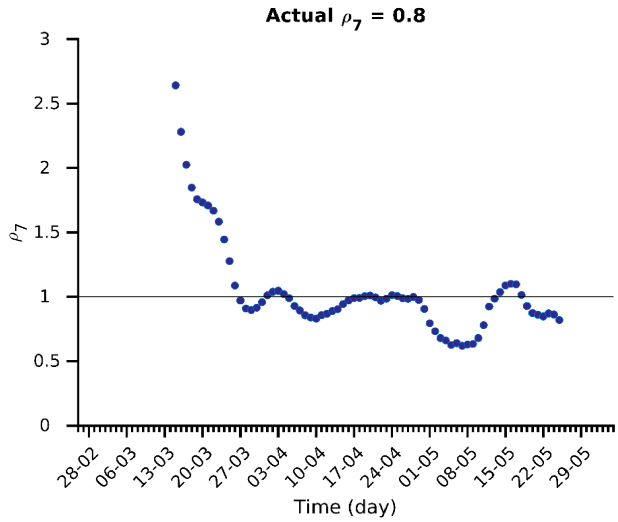


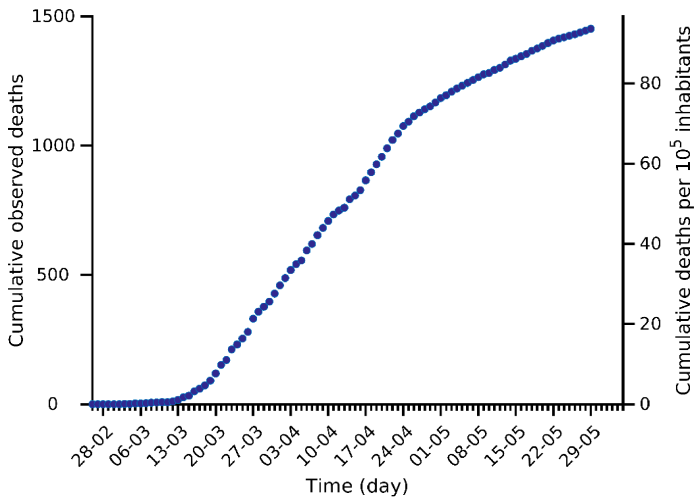


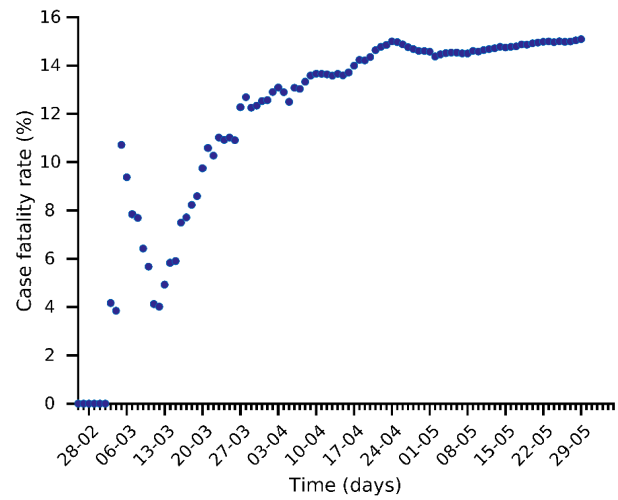




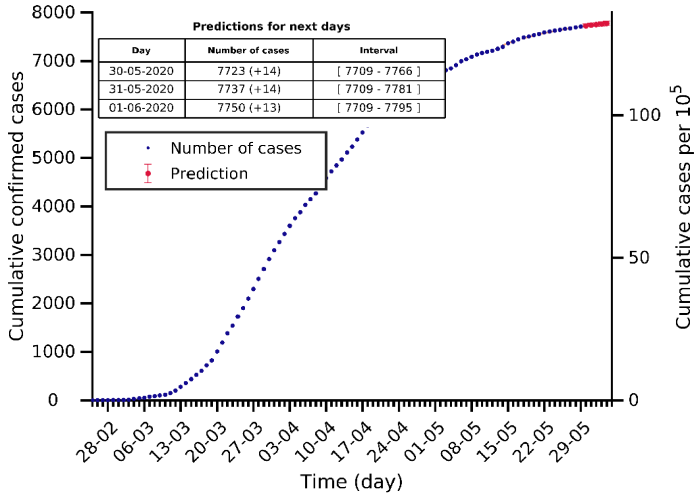


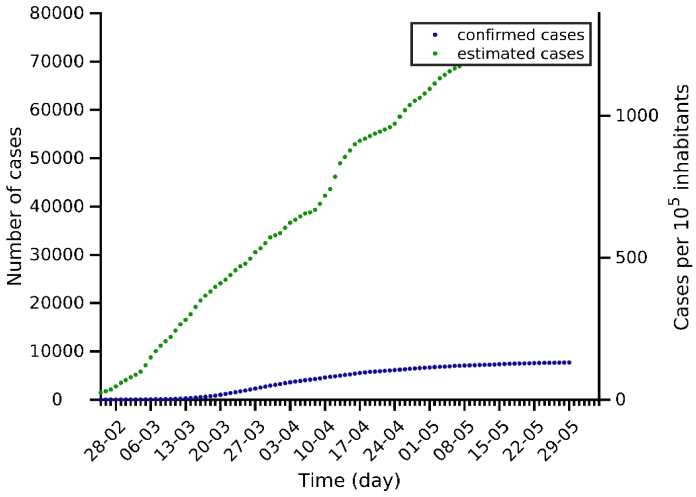


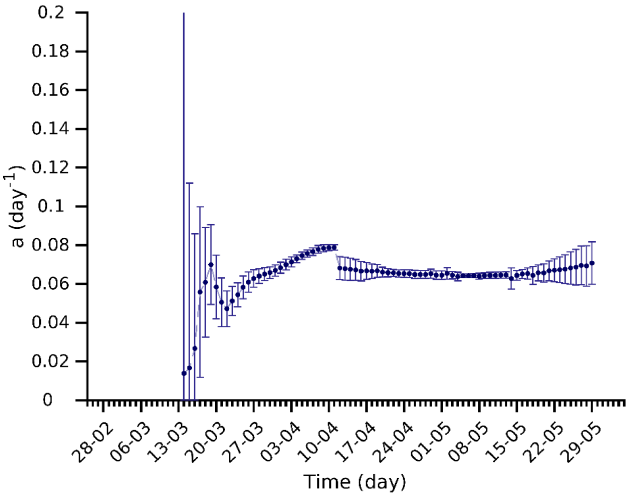


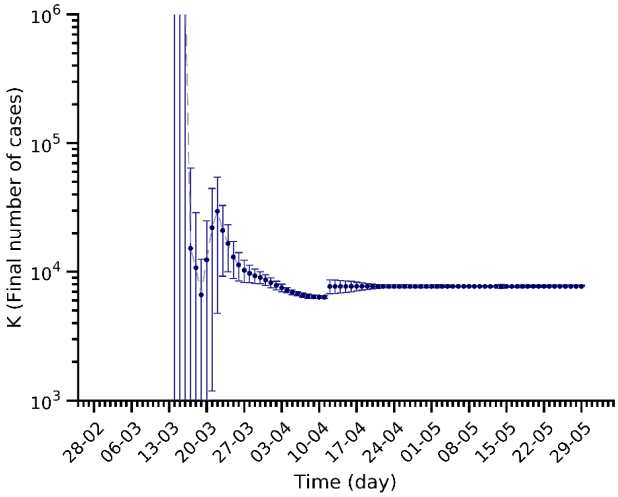


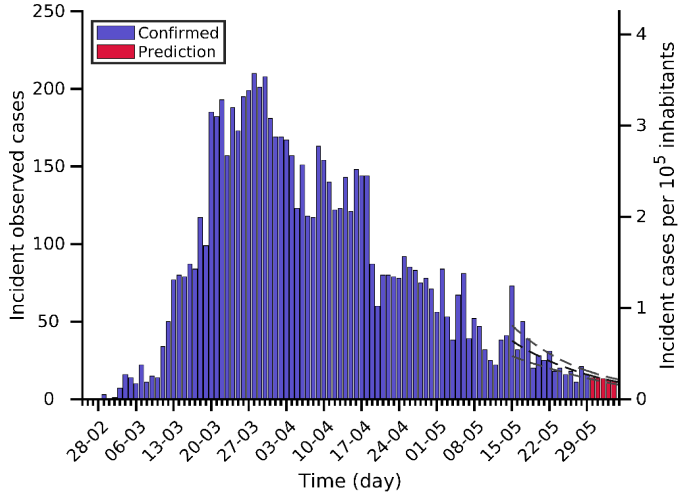
Lazio 29-05-2020. Population: 5.9M. Current cumulated incidence: 131/10⁵

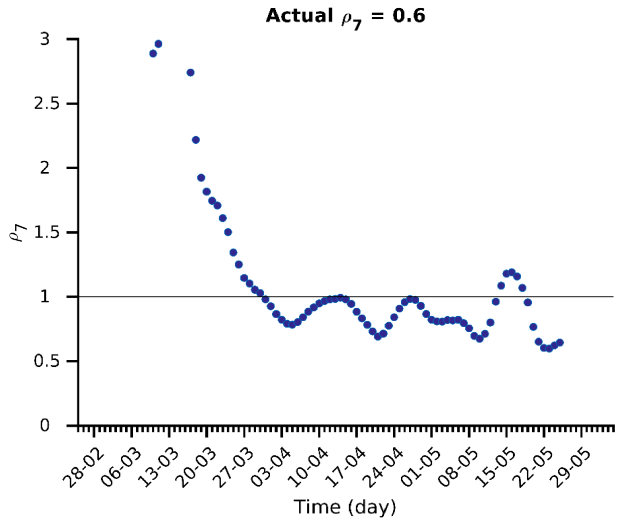


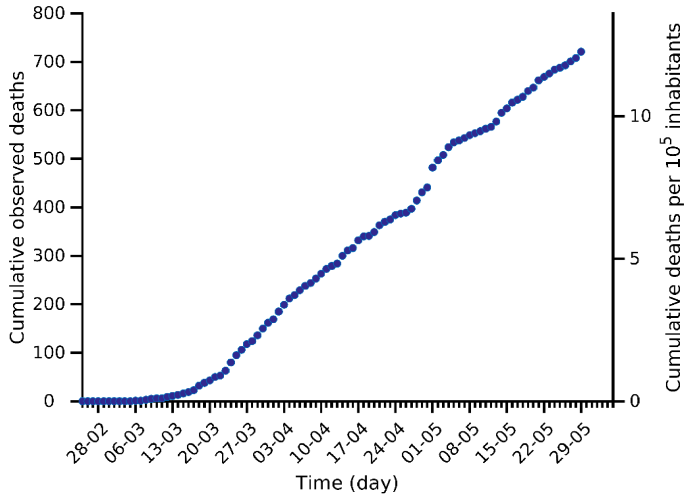


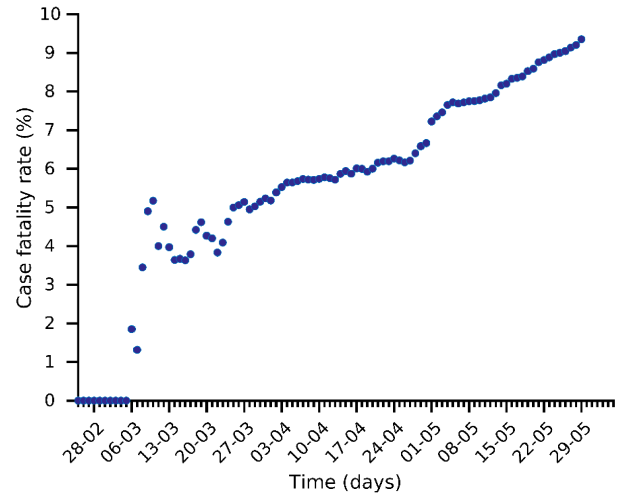




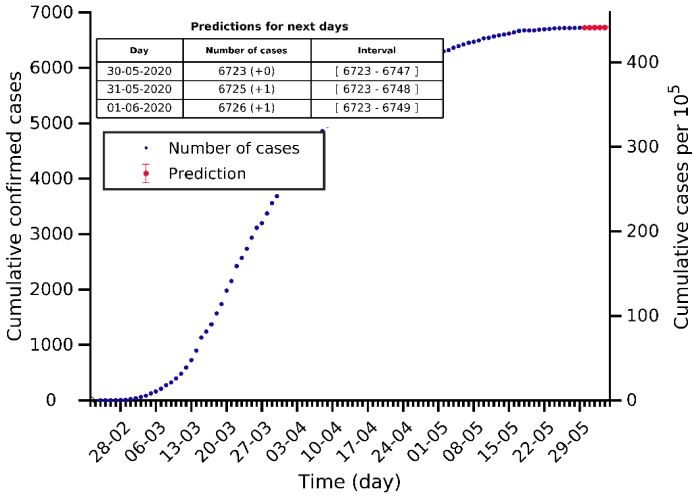


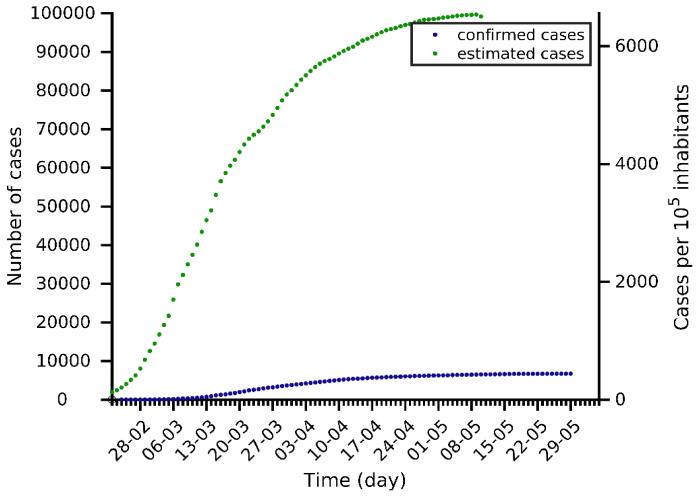


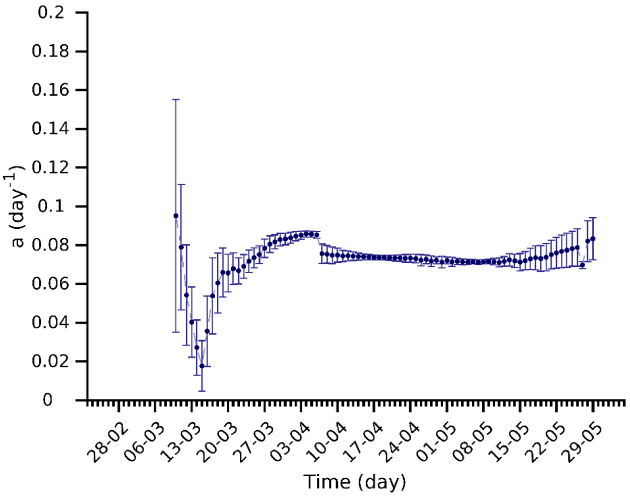


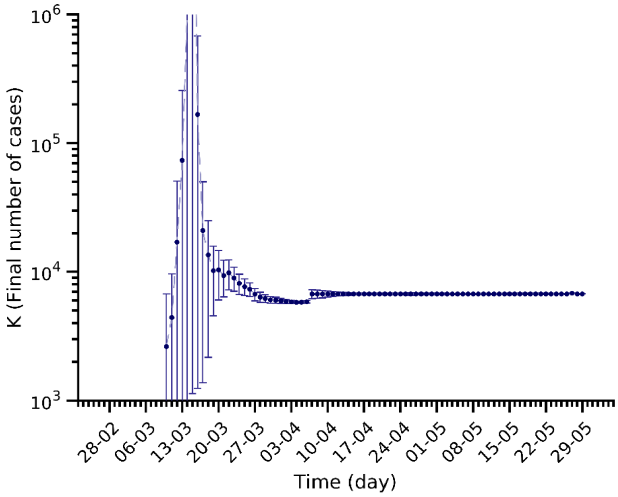


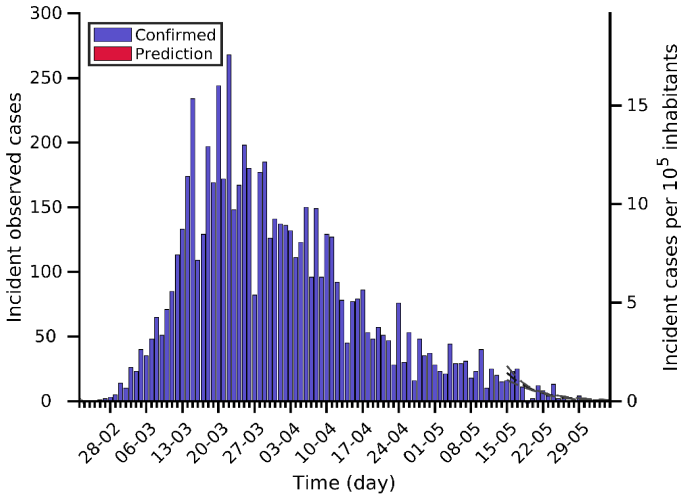
Marche 29-05-2020. Population: 1.5M. Current cumulated incidence: 441/10⁵

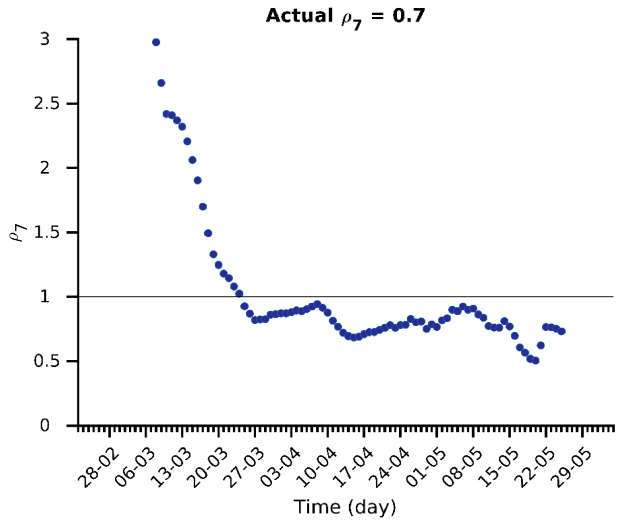


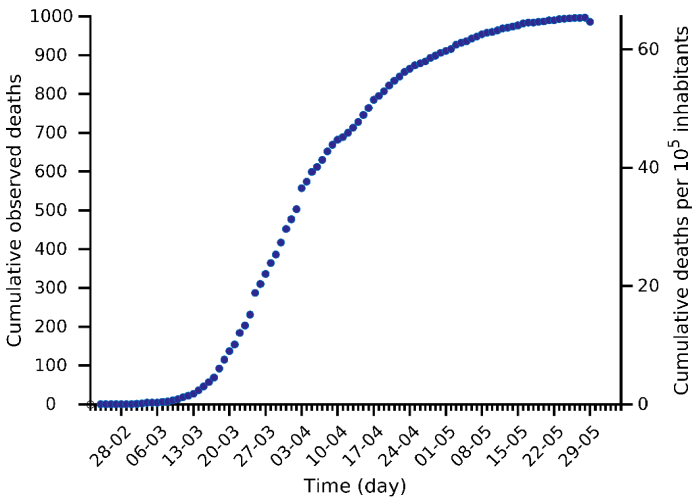


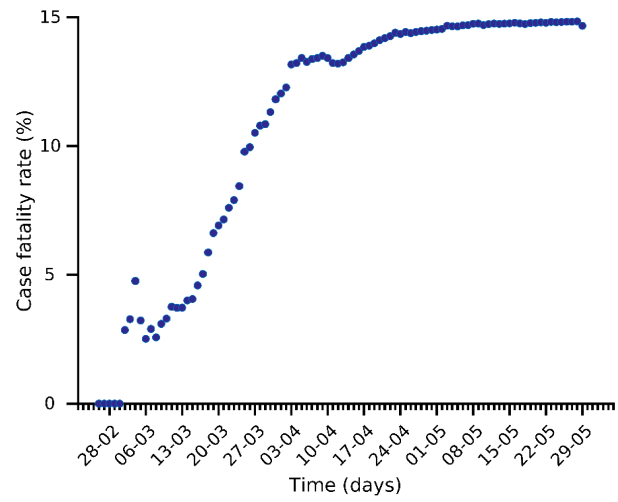




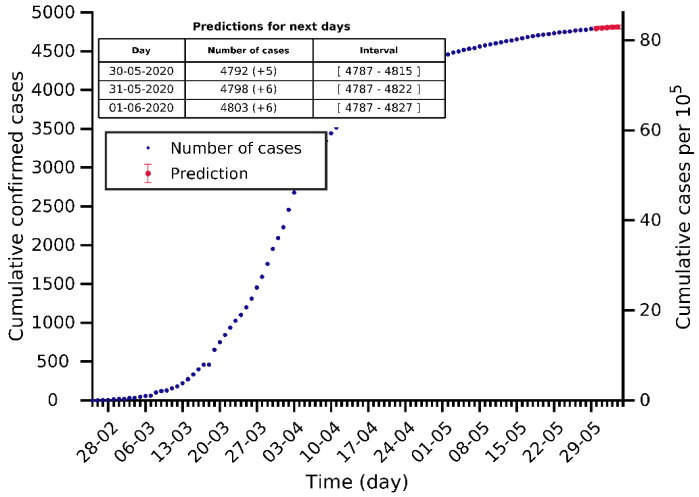


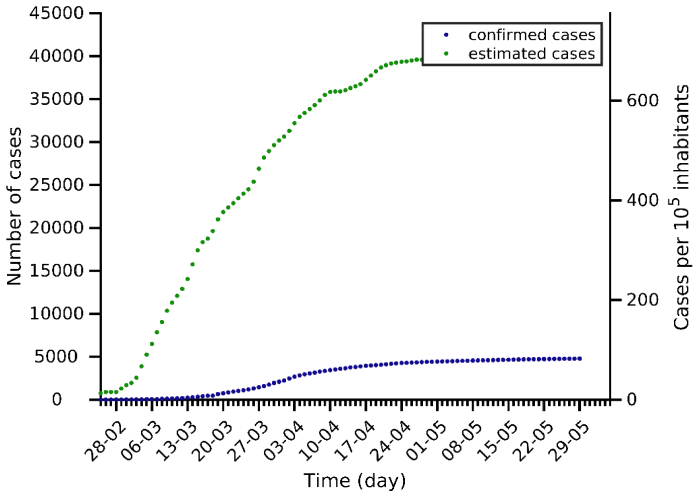


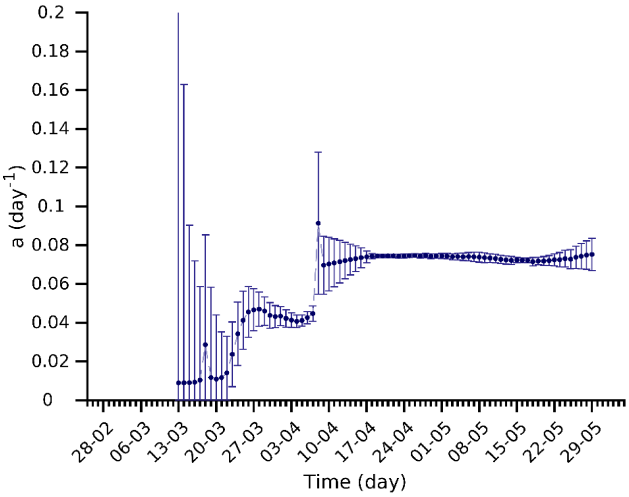


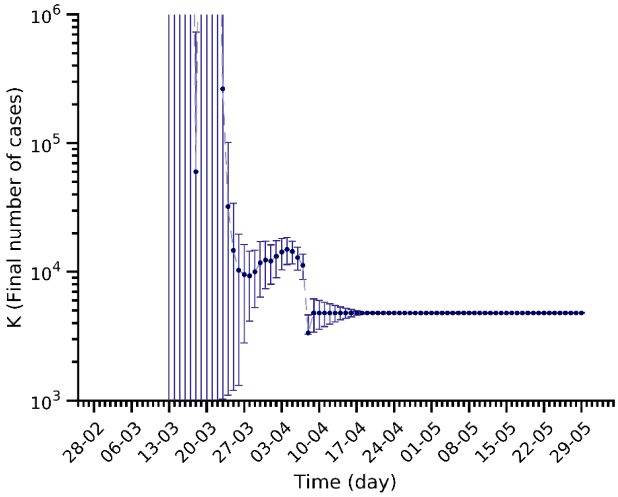


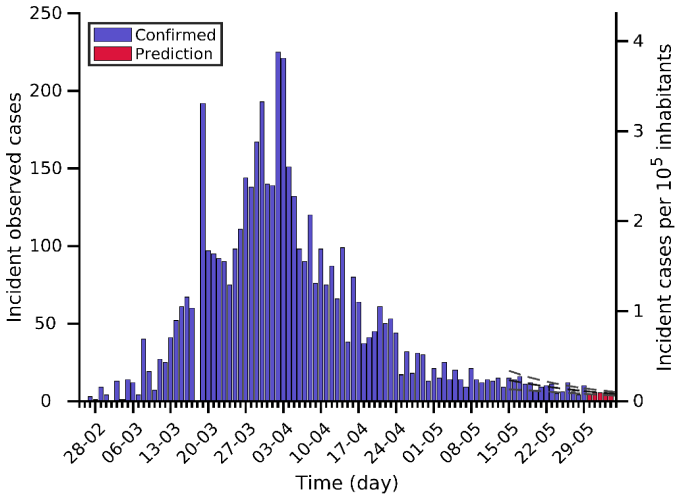
Campania 29-05-2020. Population: 5.8M. Current cumulated incidence: 83/10^5

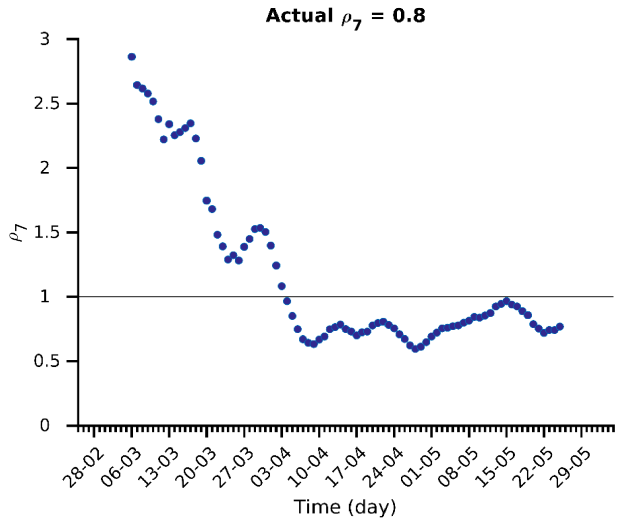


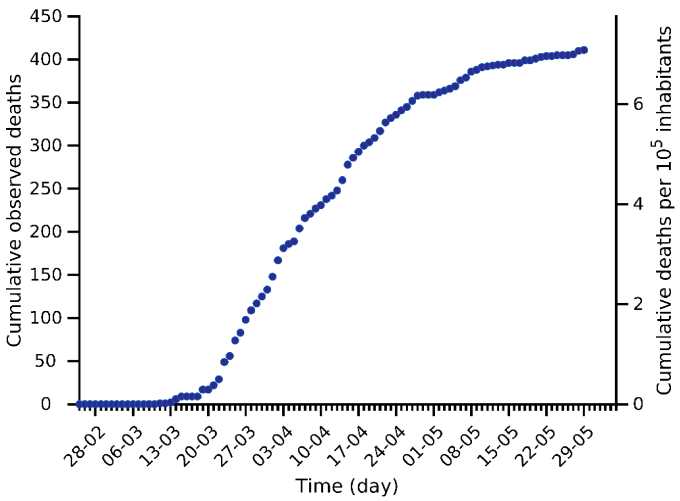


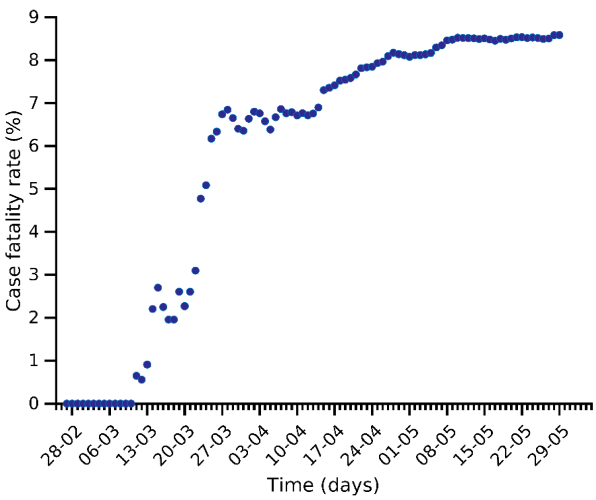




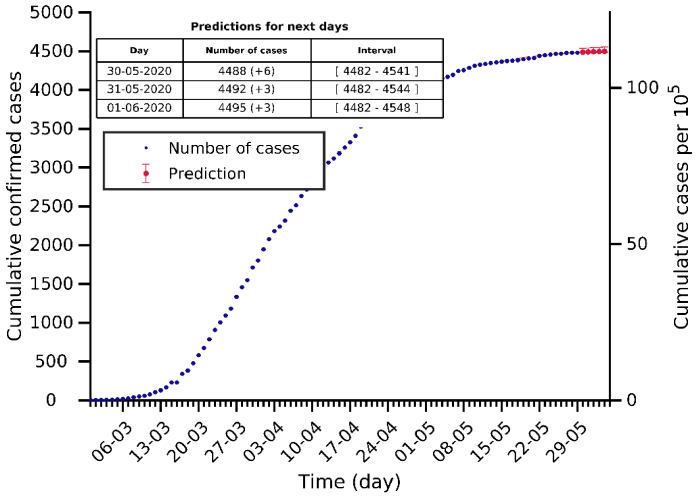


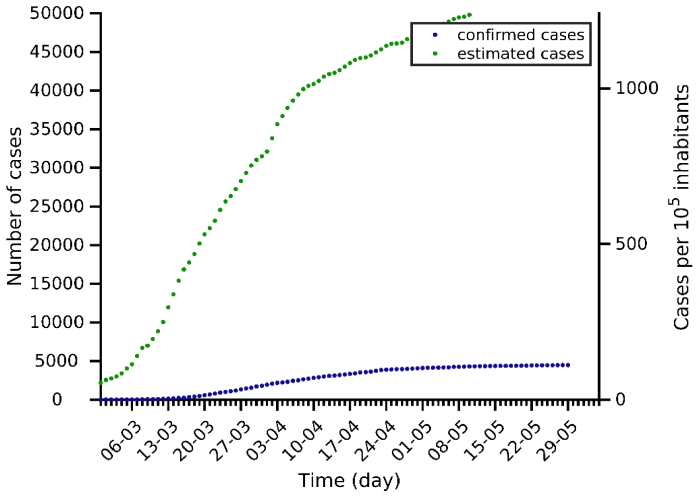


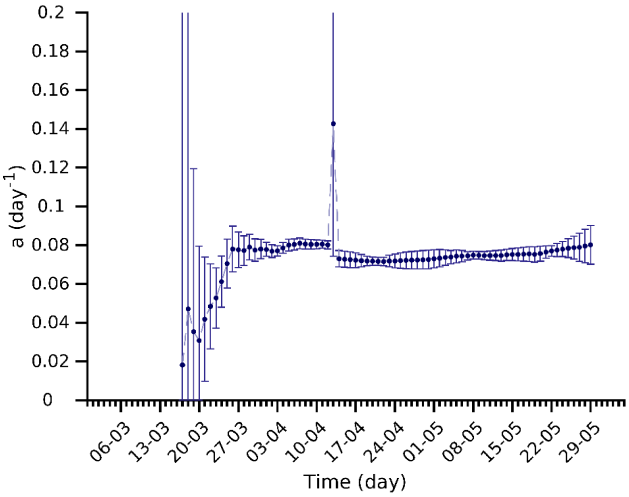


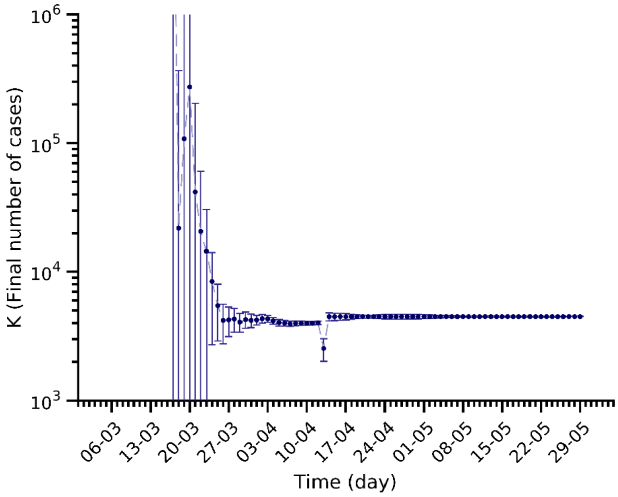


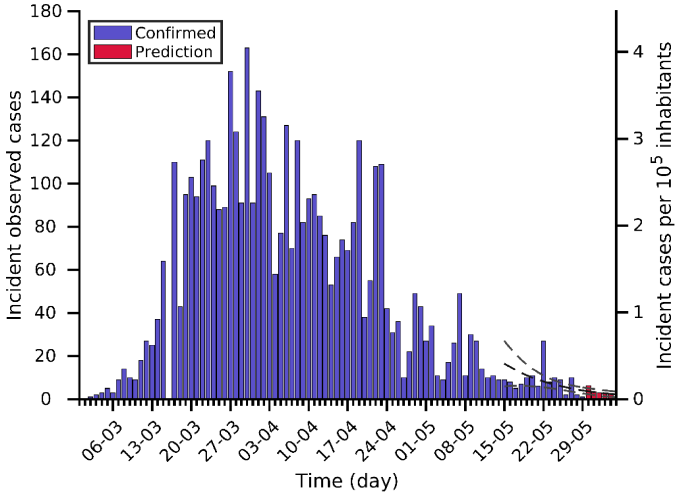
Puglia 29-05-2020. Population: 4.0M. Current cumulated incidence: 111/10^5

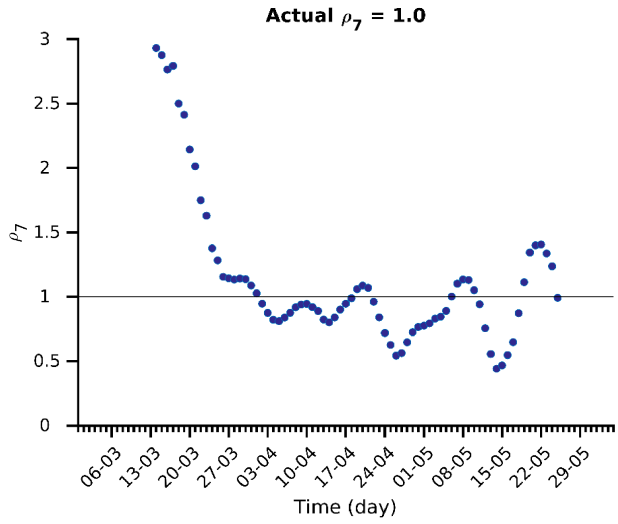


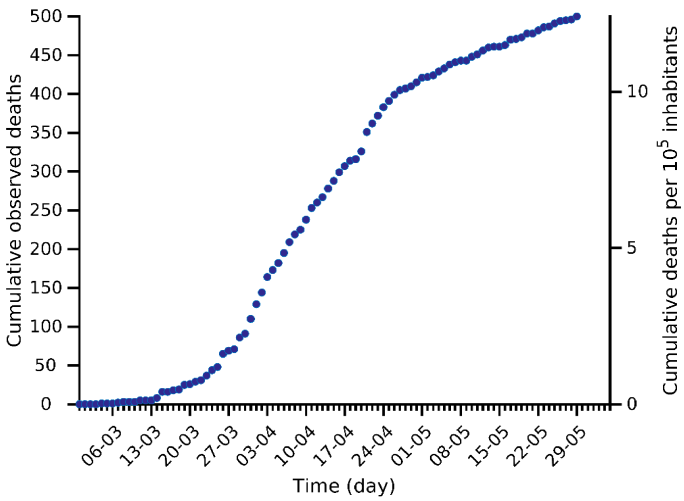


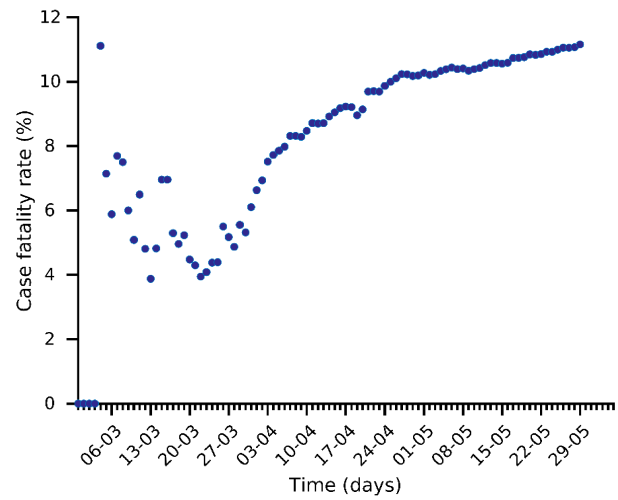




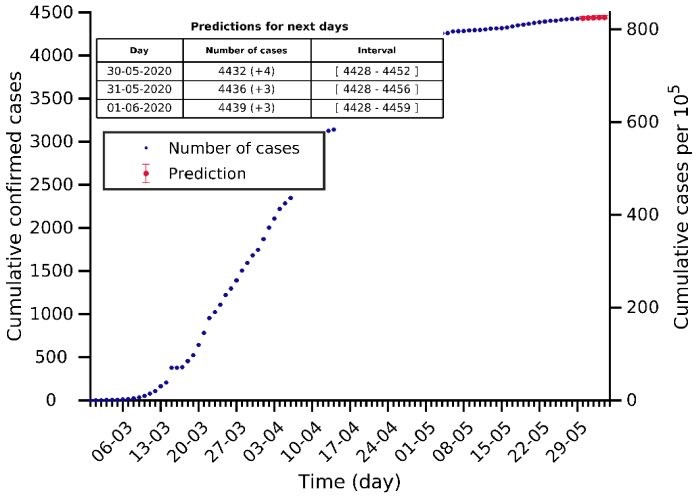


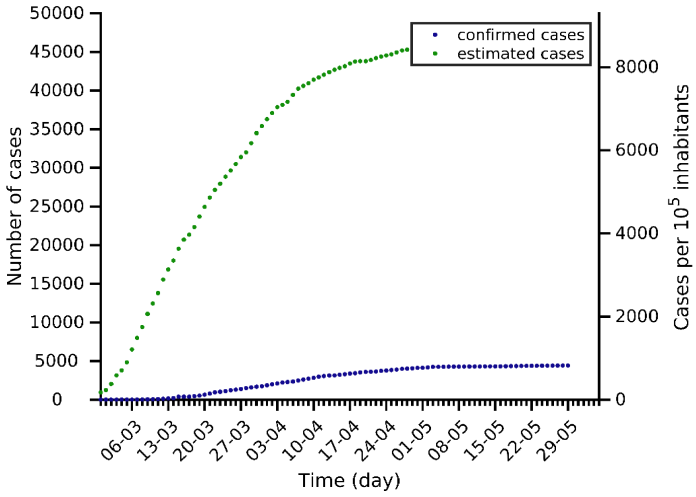


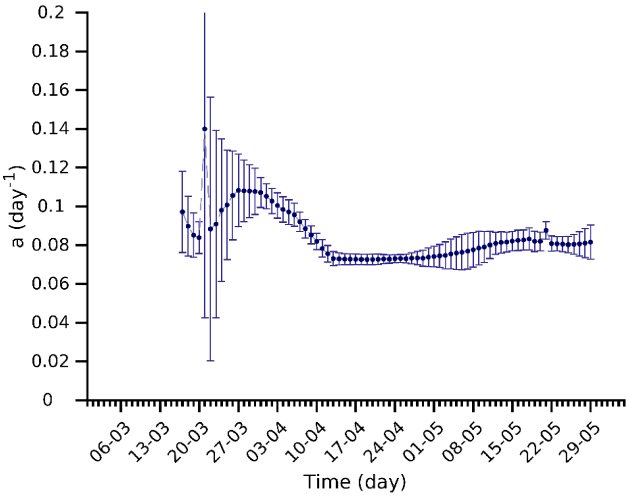


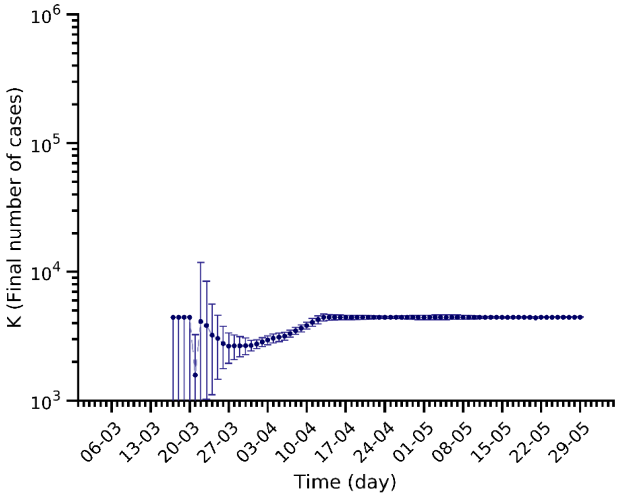


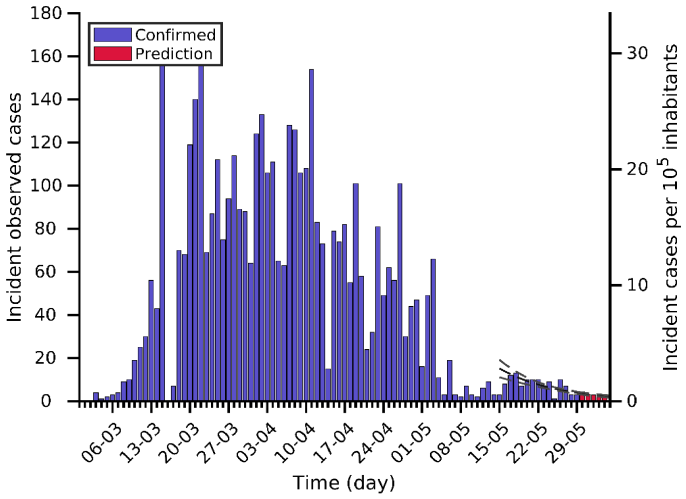
Trento 29-05-2020. Population: 0.5M. Current cumulated incidence: $823/10^5$

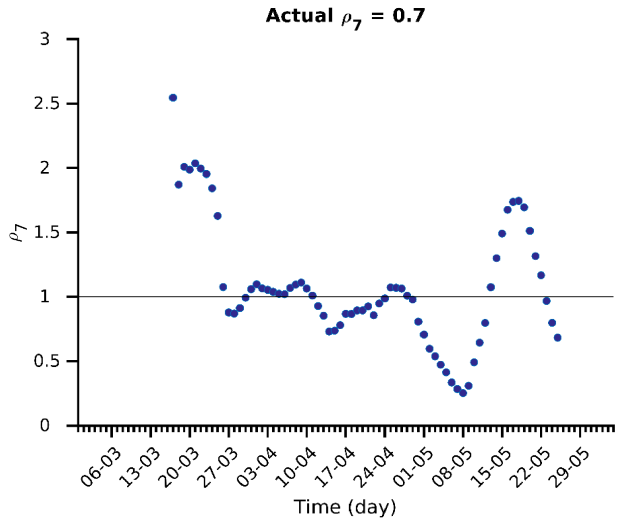


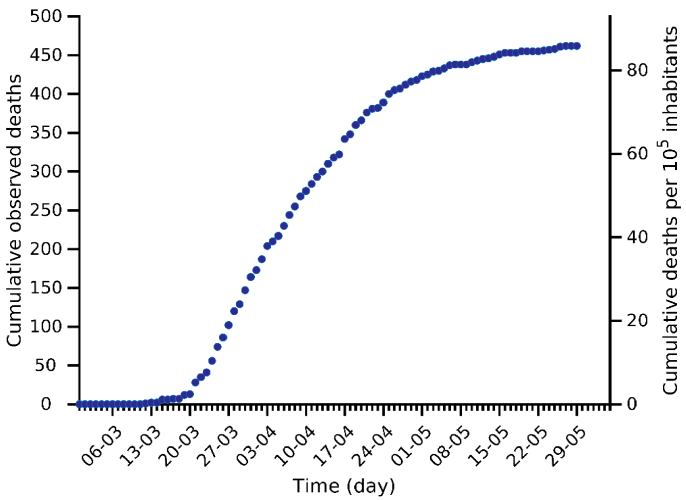


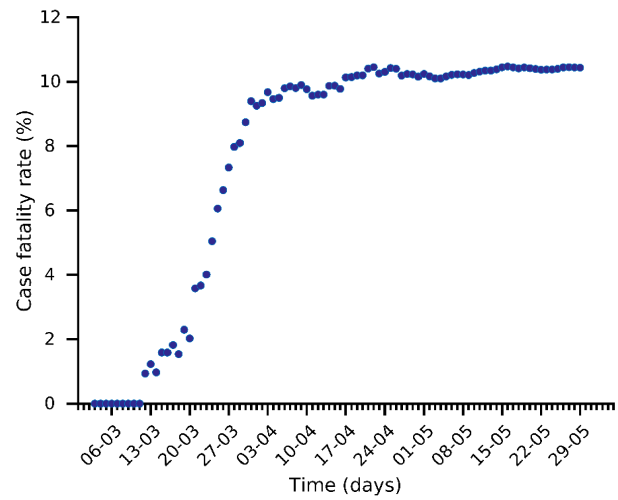




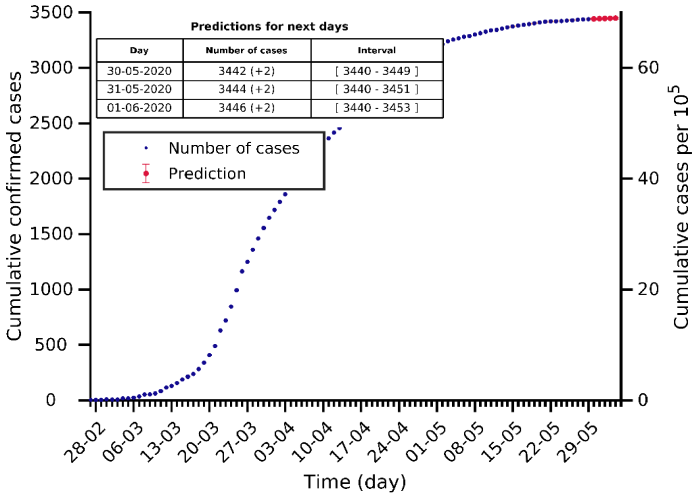


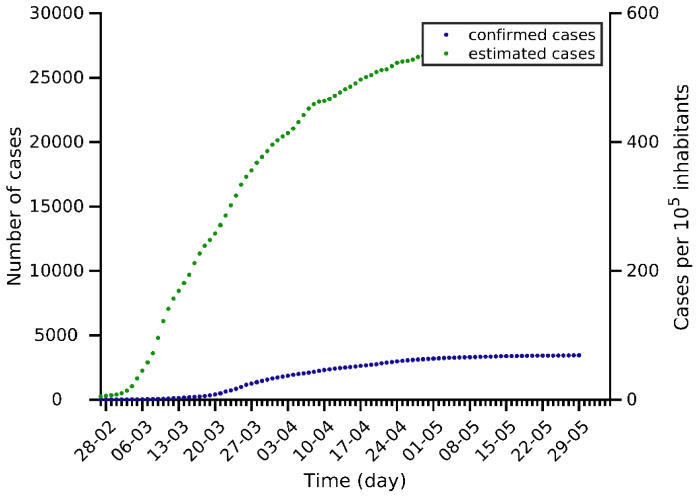


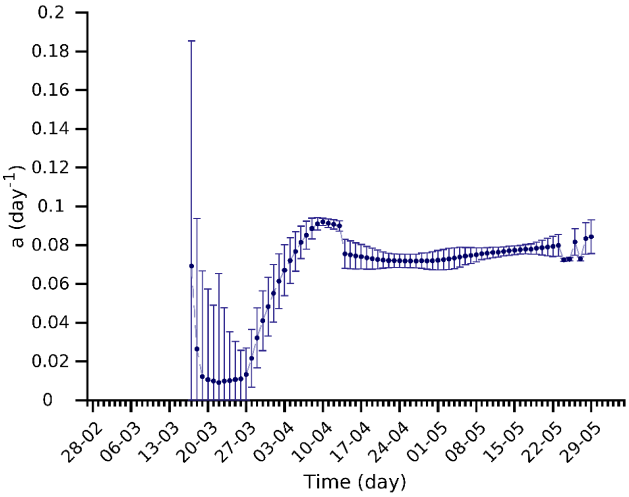


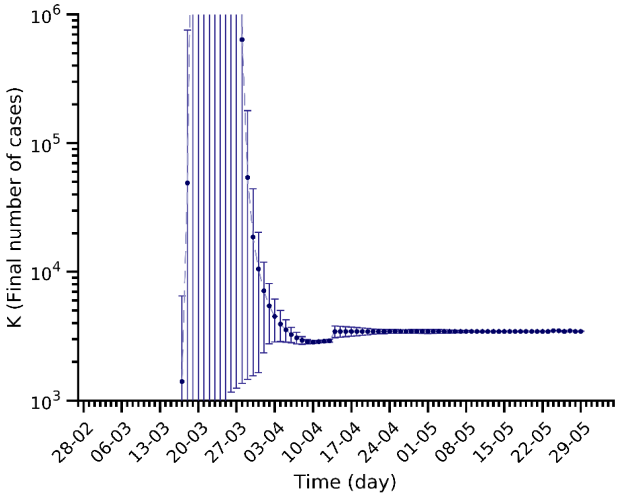


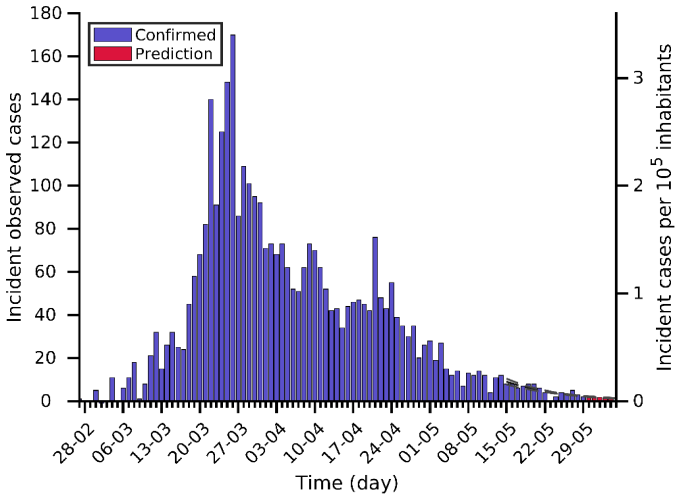
Sicilia 29-05-2020. Population: 5.0M. Current cumulated incidence: $69/10^5$

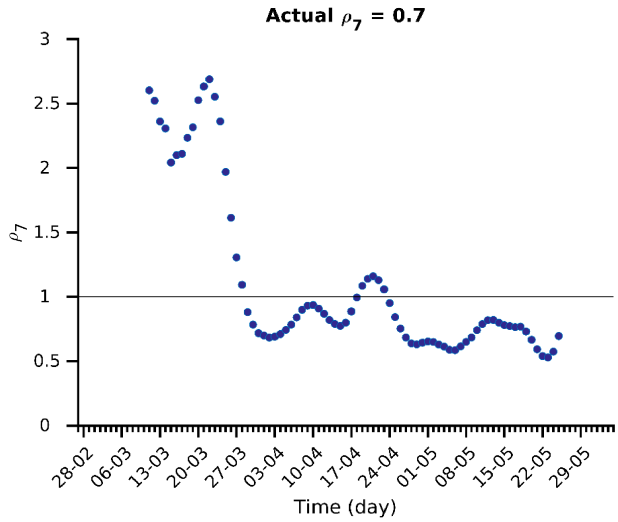


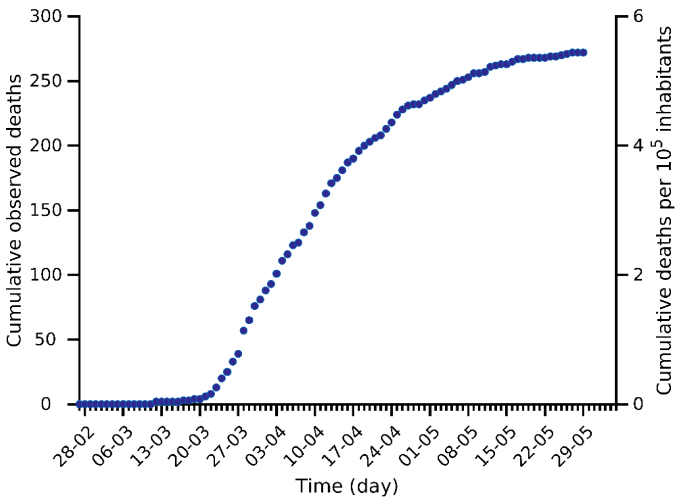


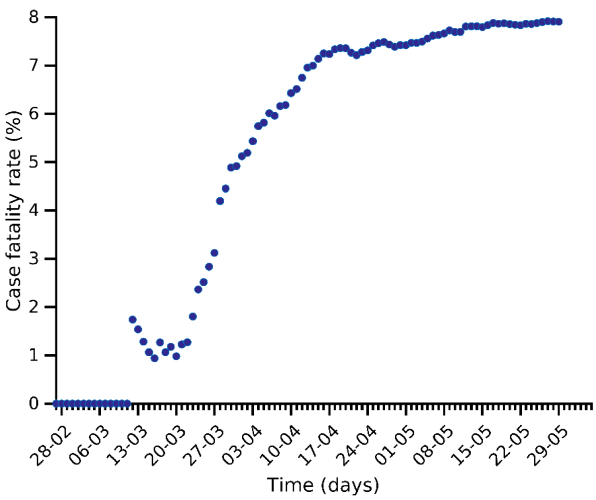




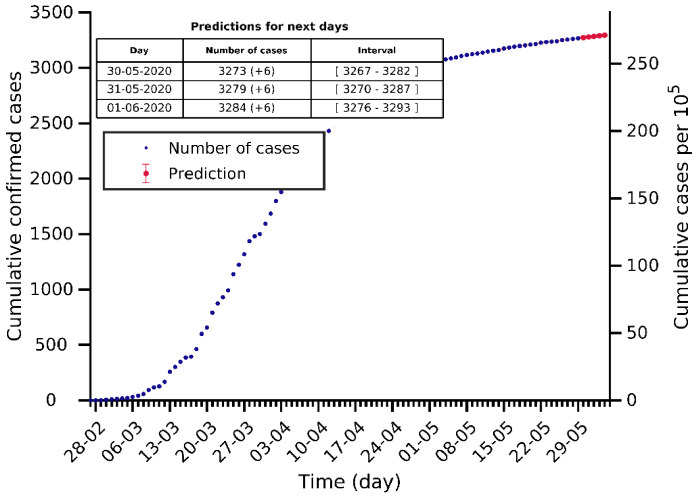


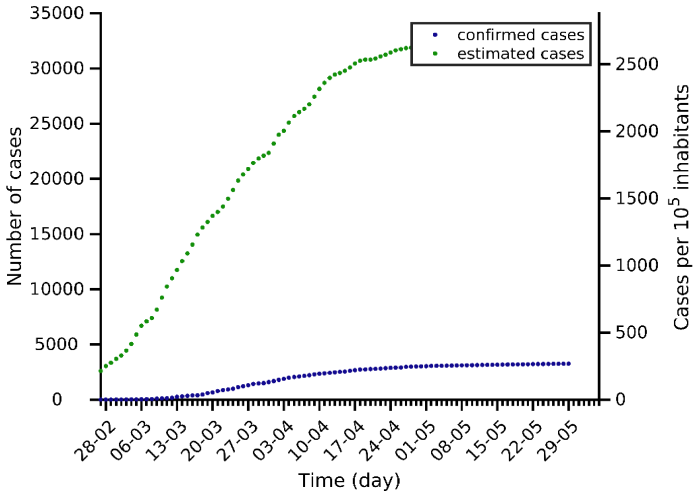


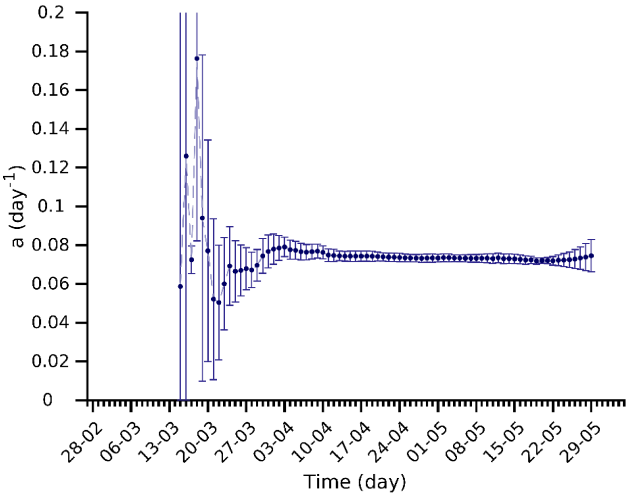


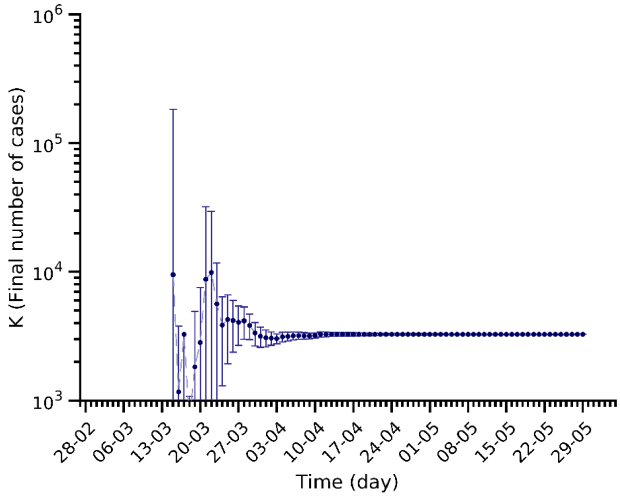


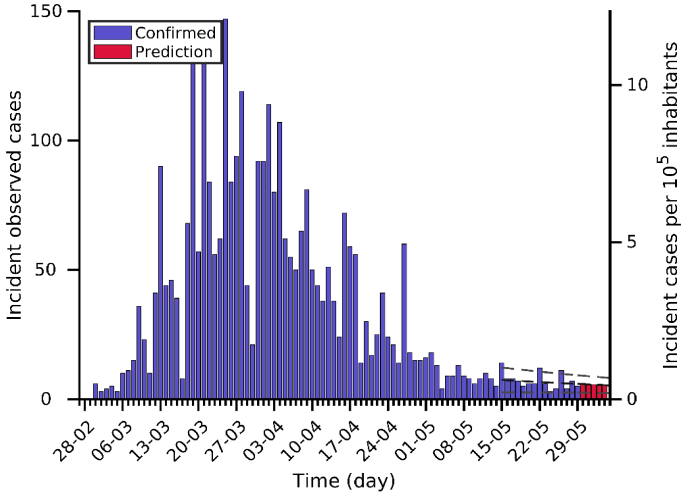
Friuli Venezia Giulia 29-05-2020. Population: 1.2M. Current cumulated incidence: 2

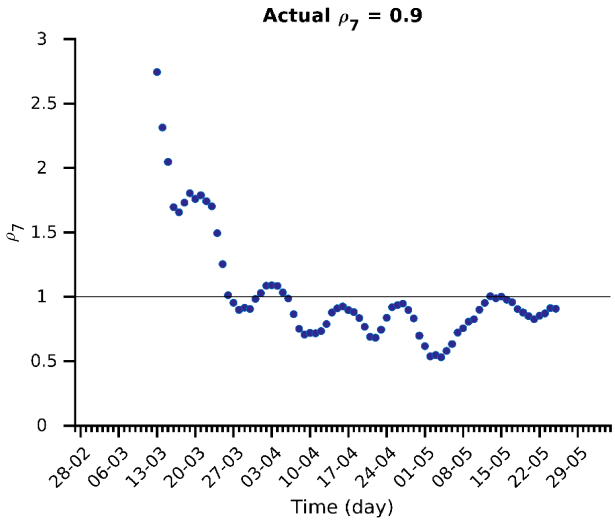


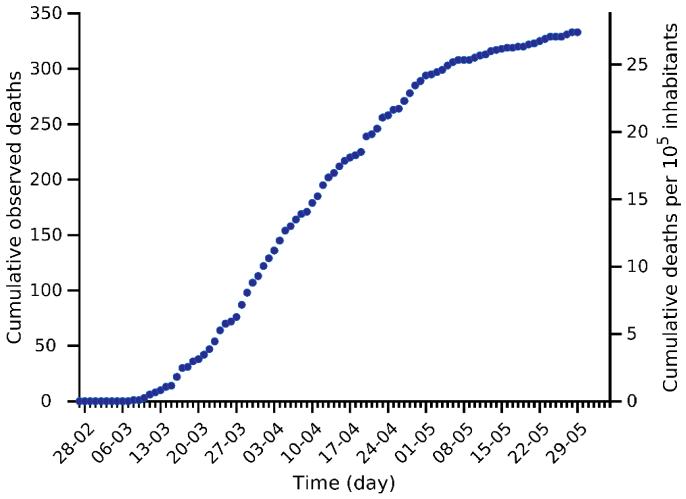


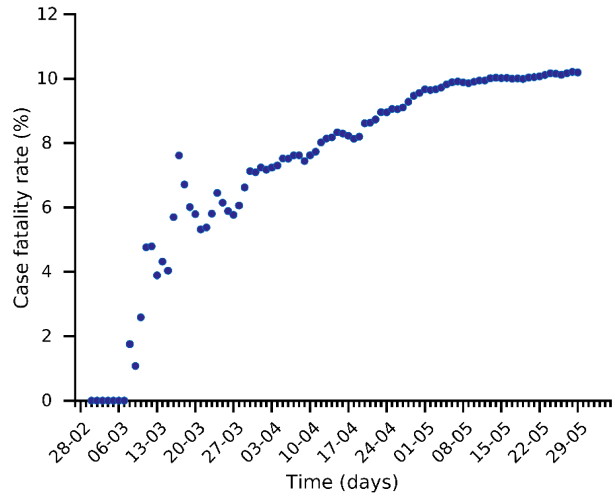




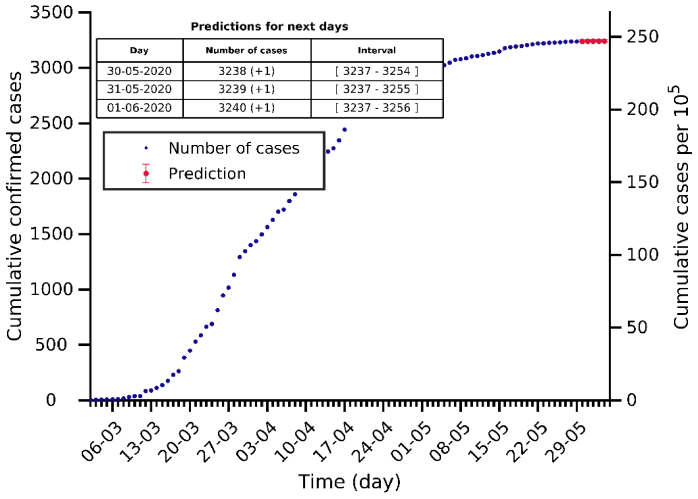


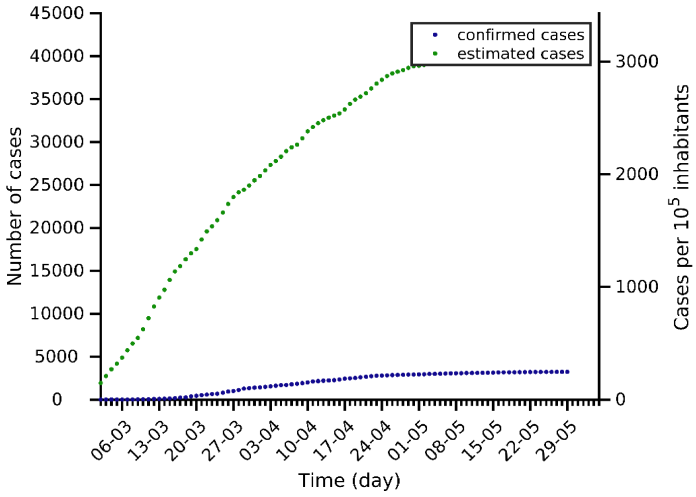


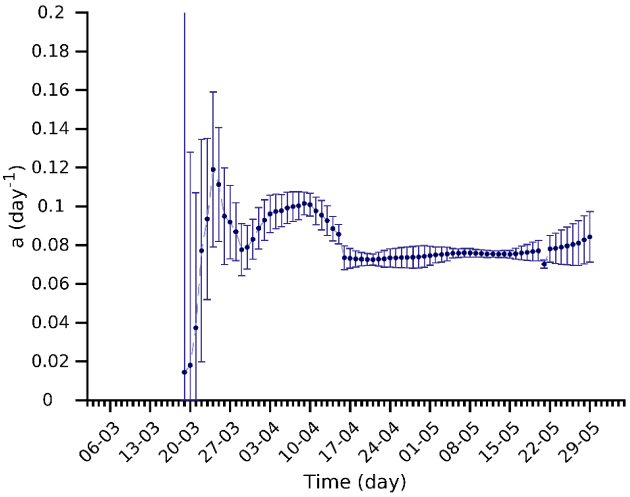


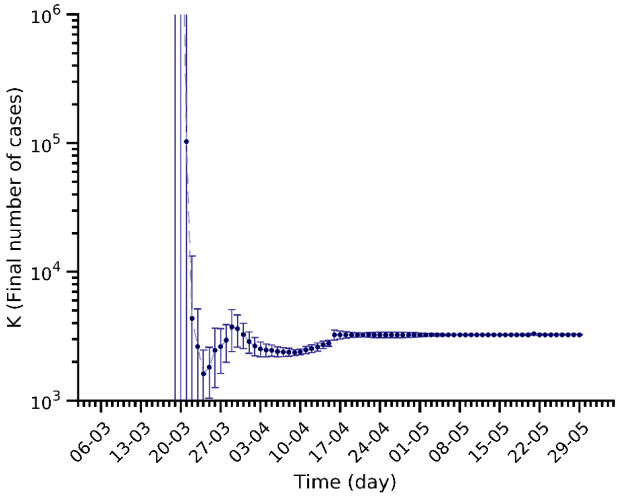


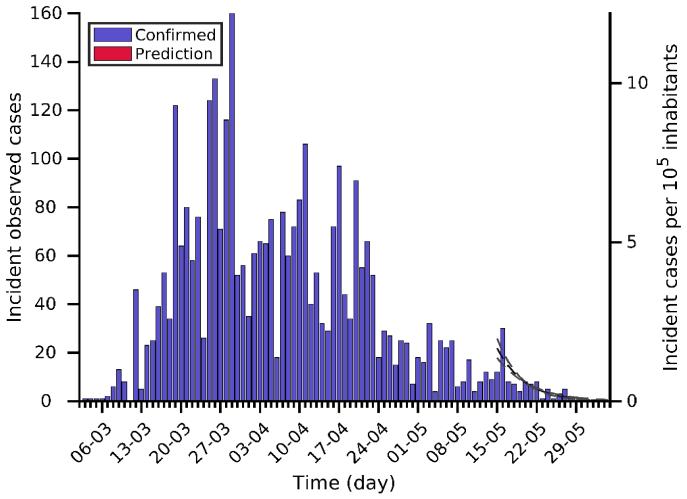
Abruzzo 29-05-2020. Population: 1.3M. Current cumulated incidence: 247/10^5

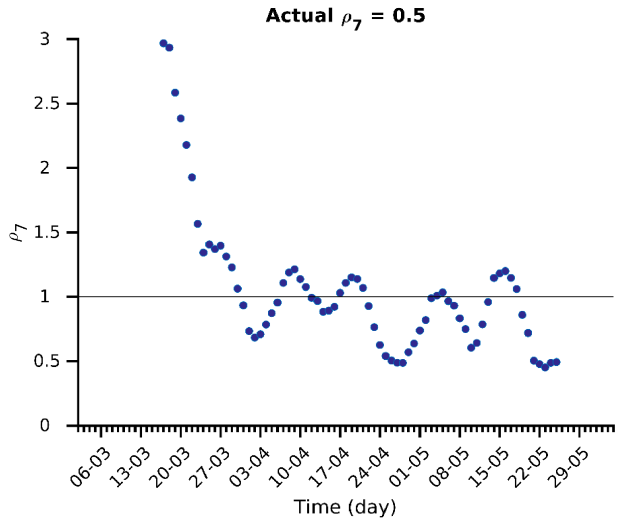


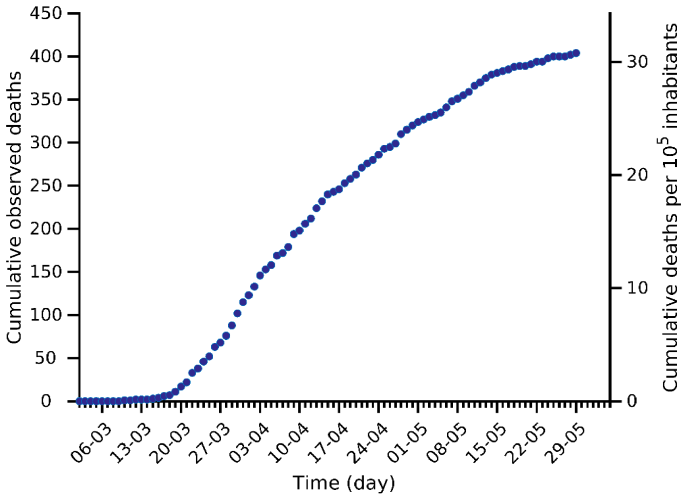


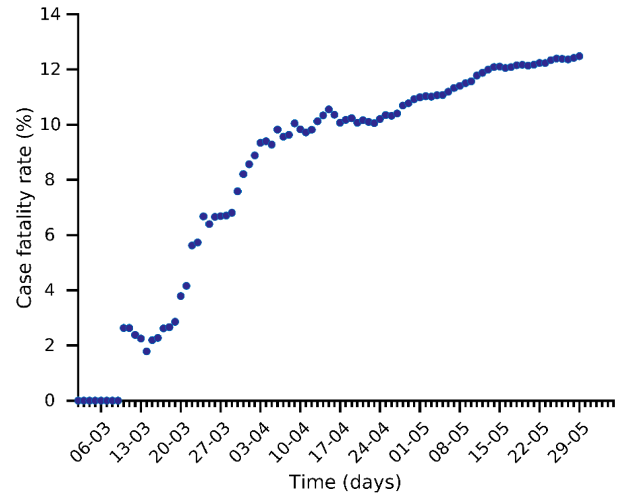




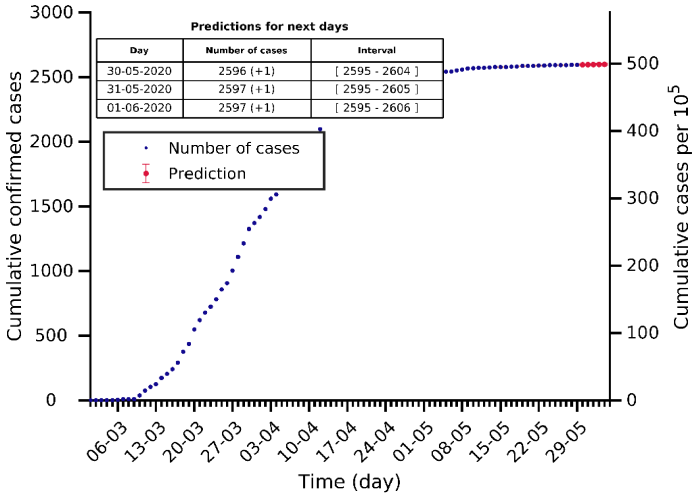


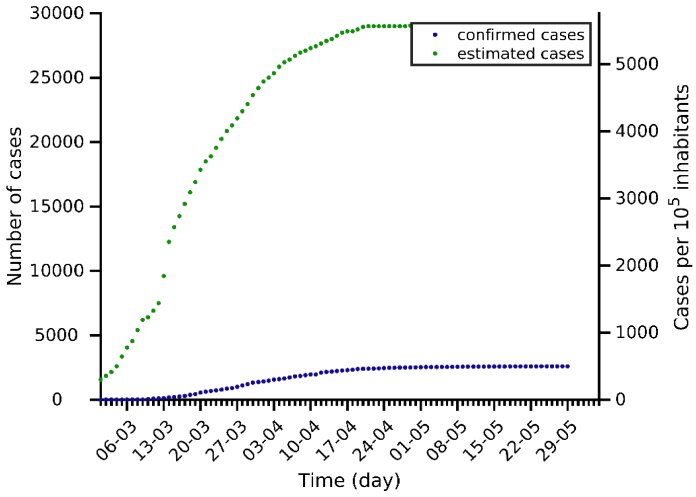


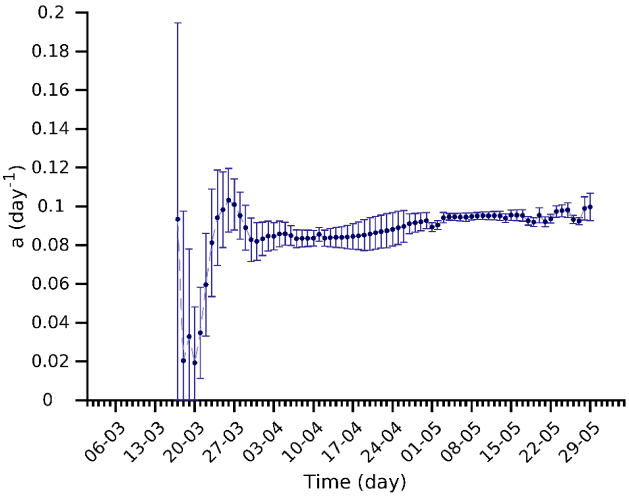


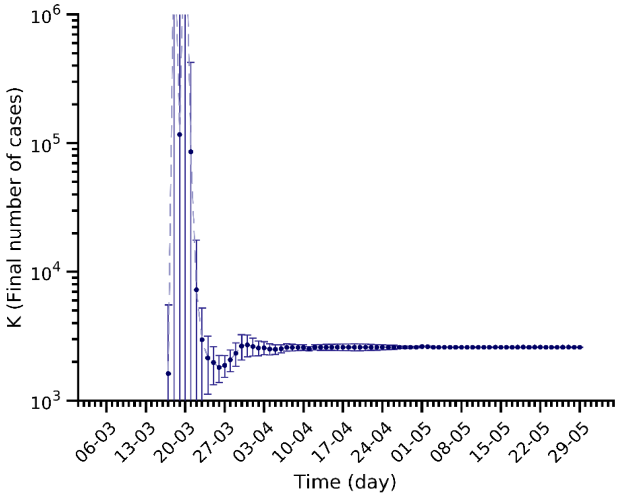


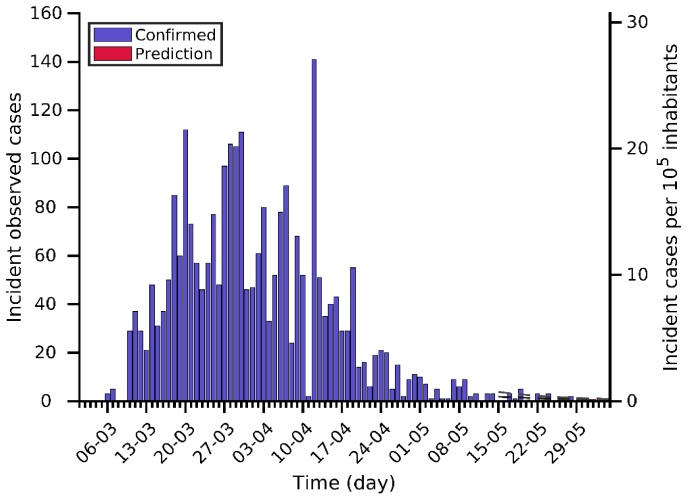
Bolzano 29-05-2020. Population: 0.5M. Current cumulated incidence: 498/10^5

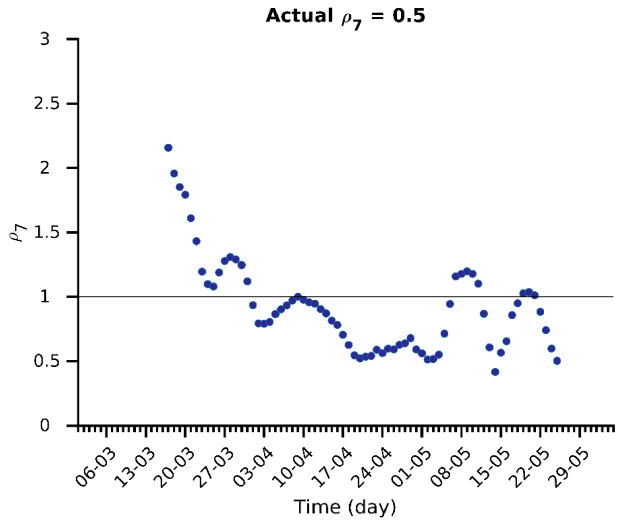


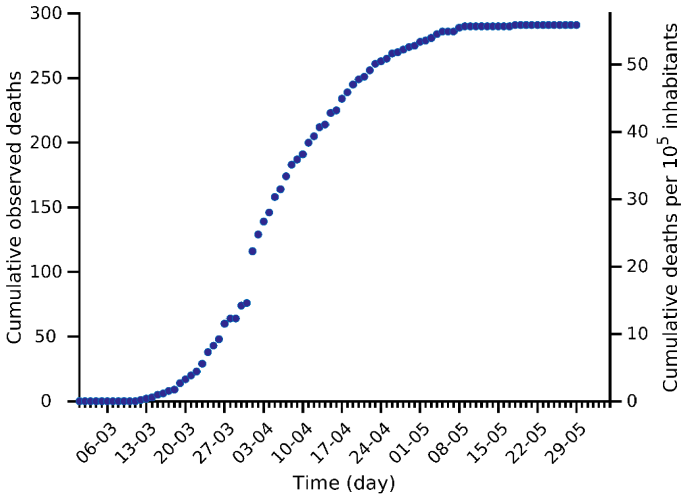


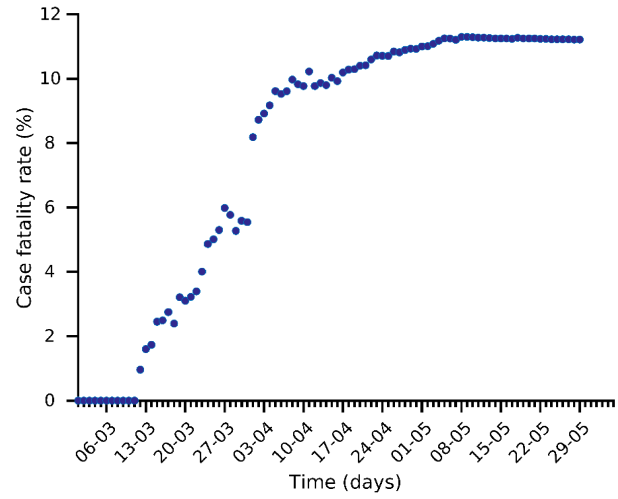




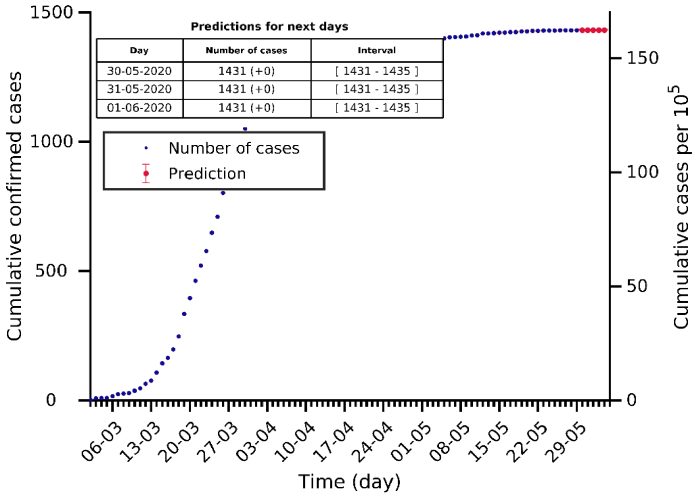


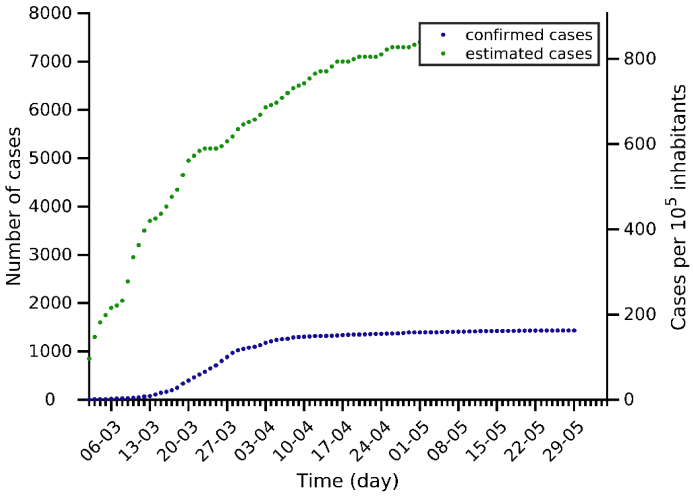


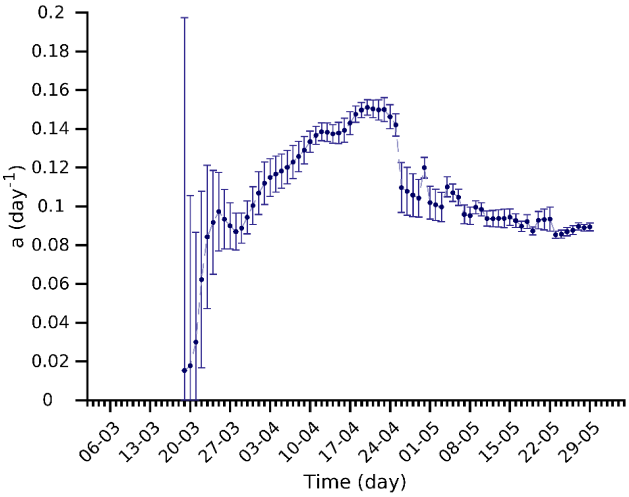


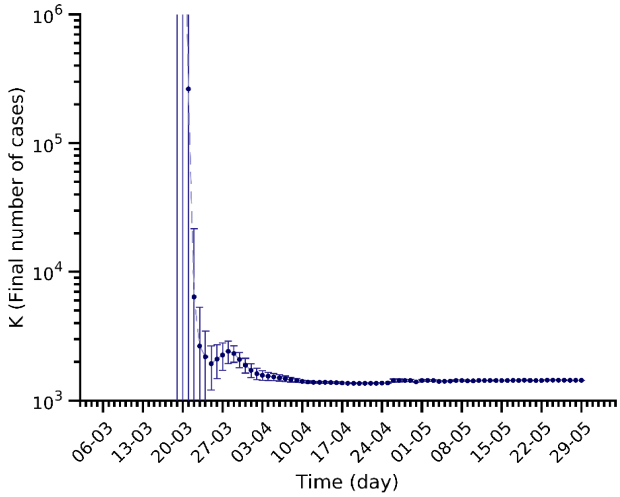


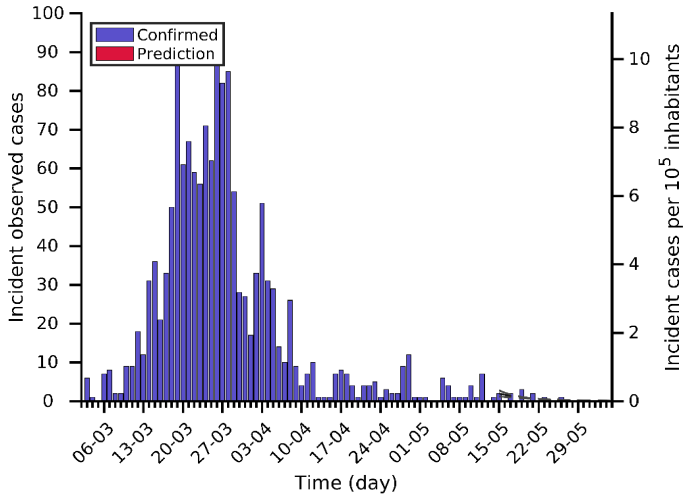
Umbria 29-05-2020. Population: 0.9M. Current cumulated incidence: 162/10^5

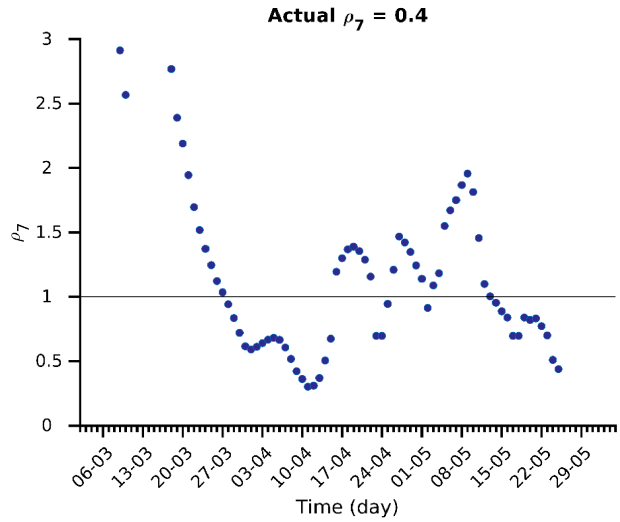


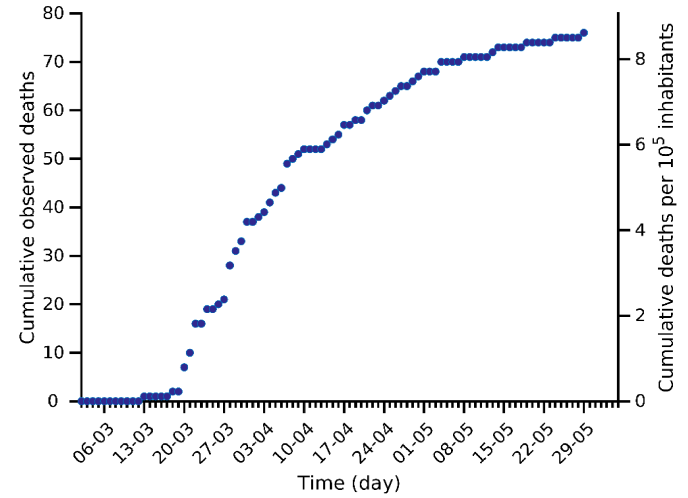


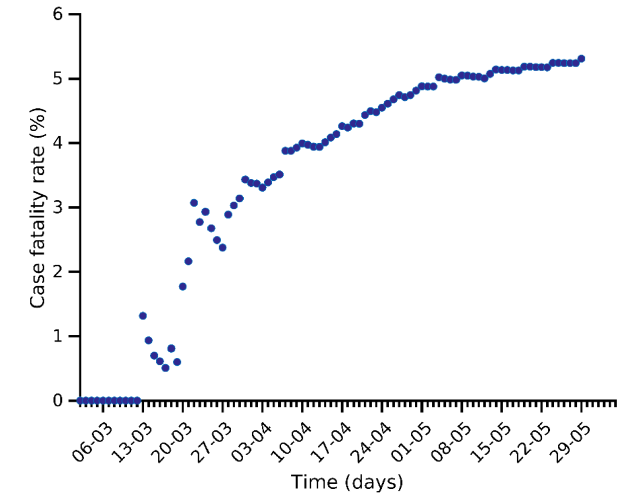








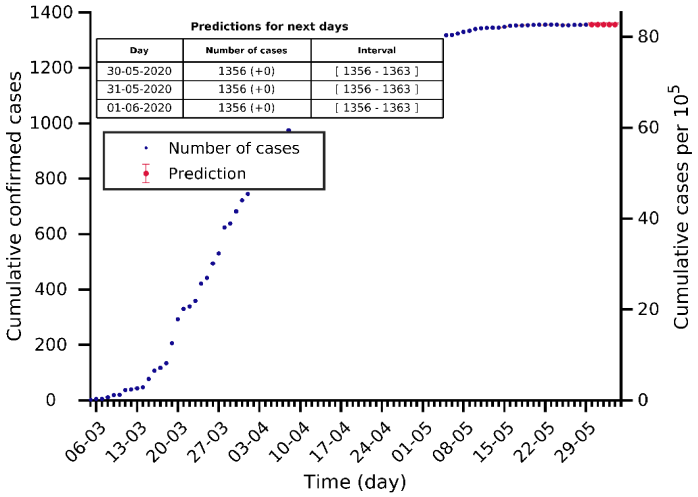


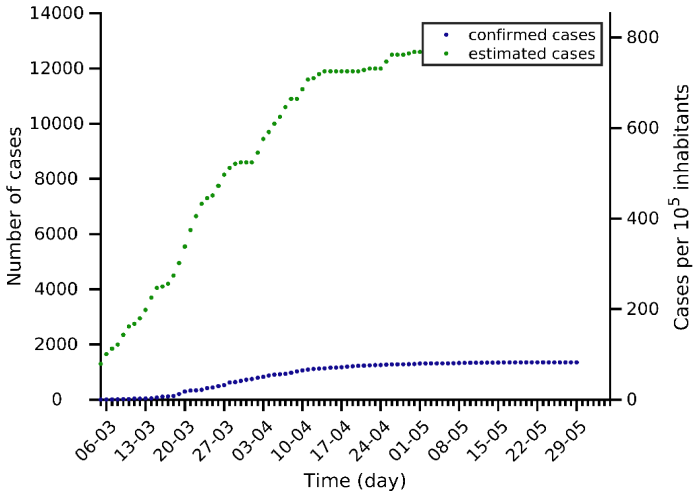


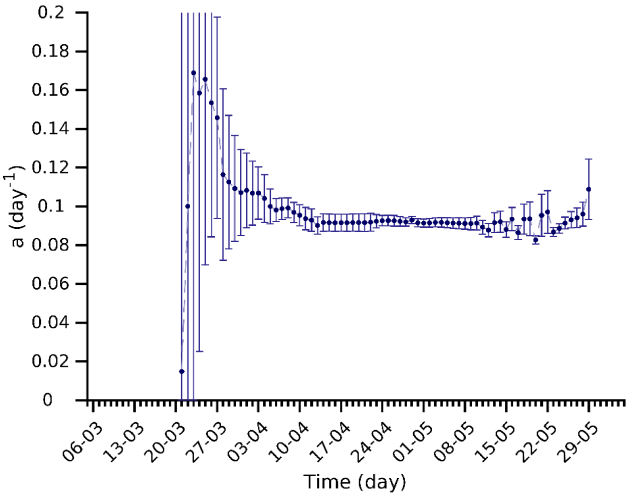
Sardegna 29-05-2020. Population: 1.6M. Current cumulated incidence: 83/10⁵

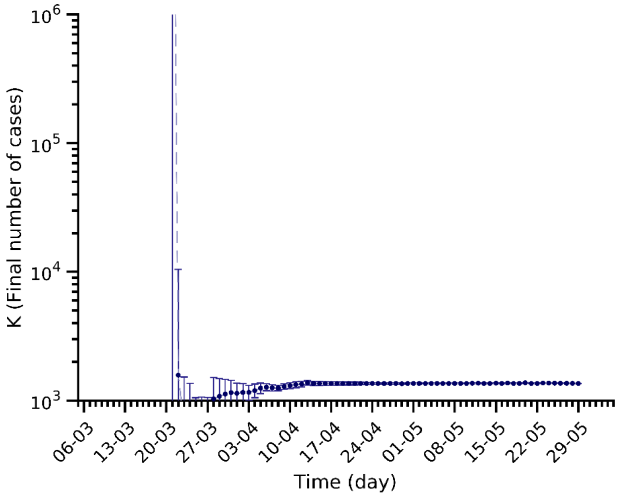
Predictions for next days

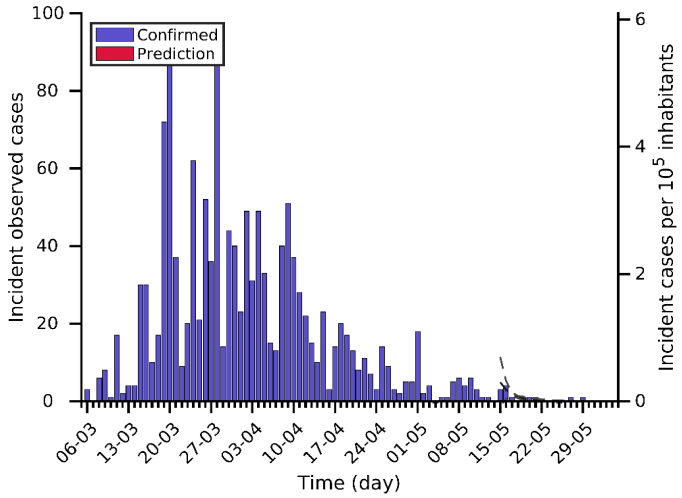
Day	Number of cases	Interval
30-05-2020	1356 (+0)	[1356 - 1363]
31-05-2020	1356 (+0)	[1356 - 1363]
01-06-2020	1356 (+0)	[1356 - 1363]

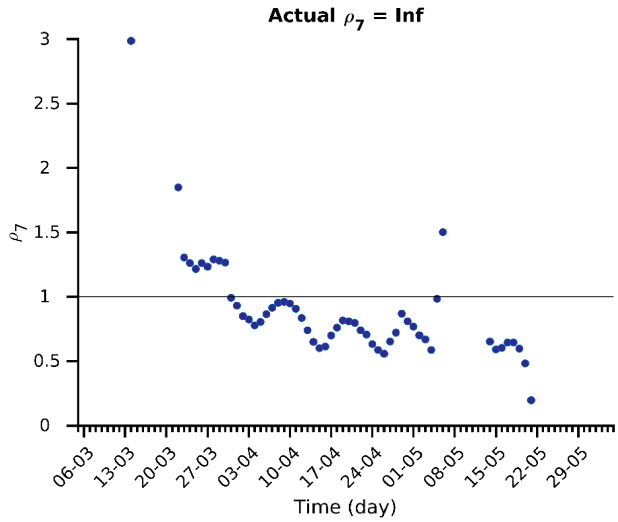


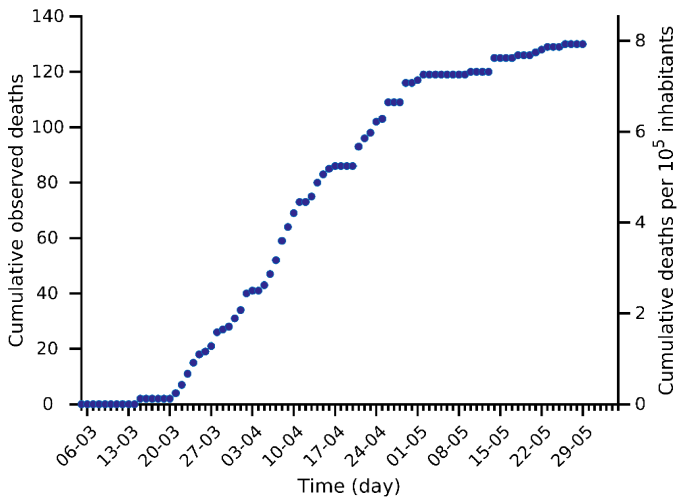


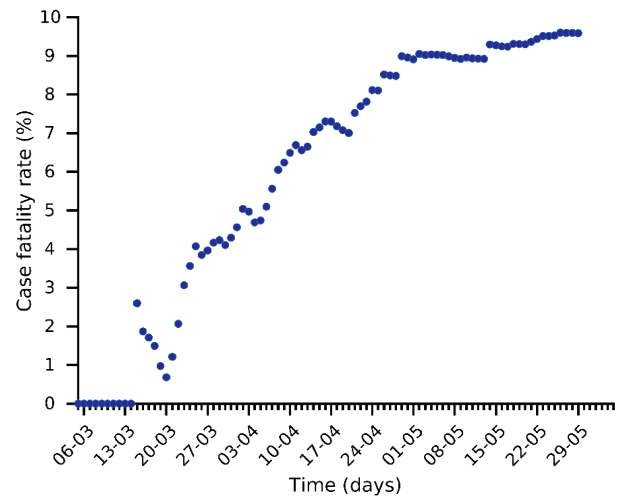




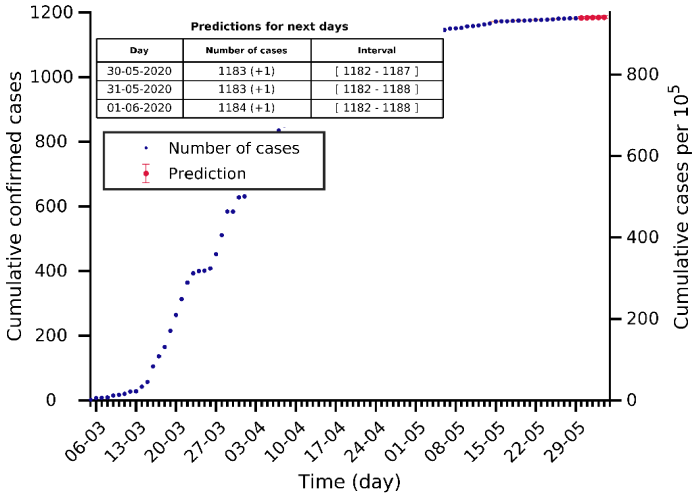


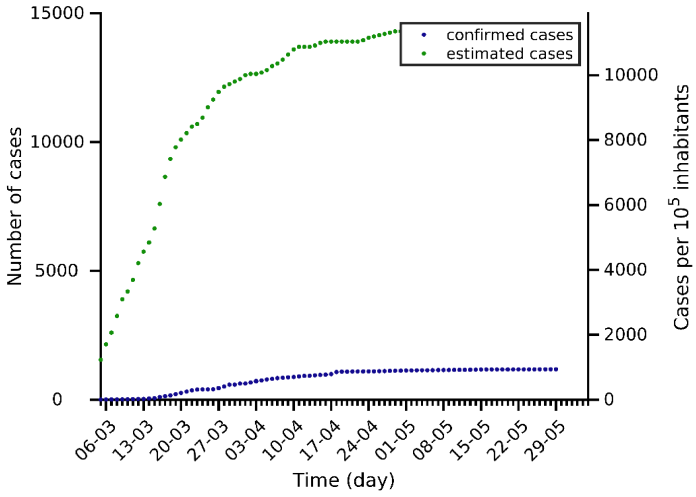


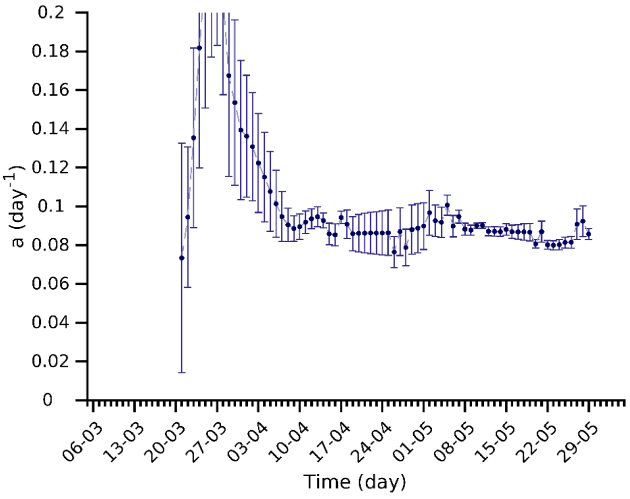


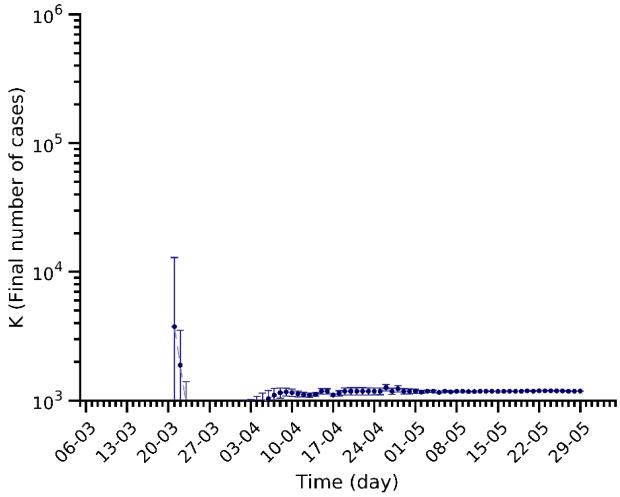


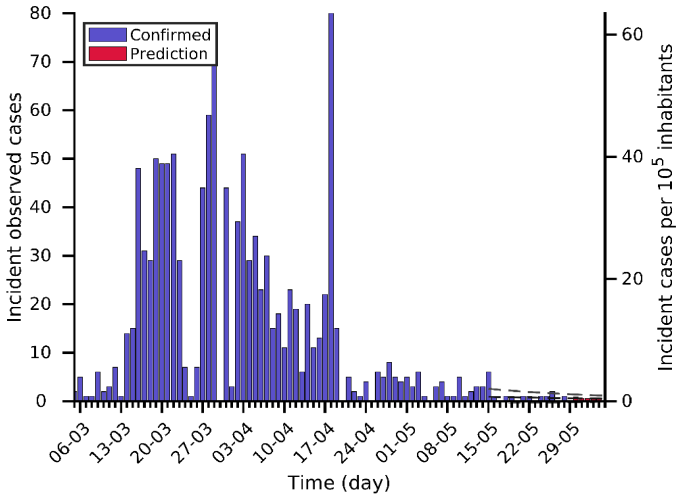
Valle dAosta 29-05-2020. Population: 0.1M. Current cumulated incidence: 938/10^5

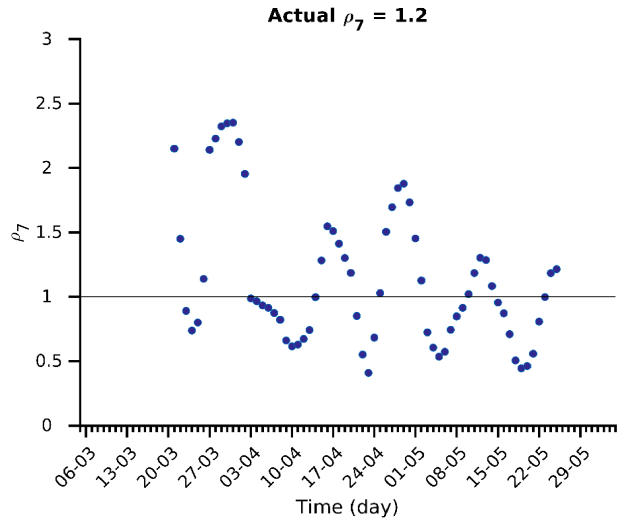


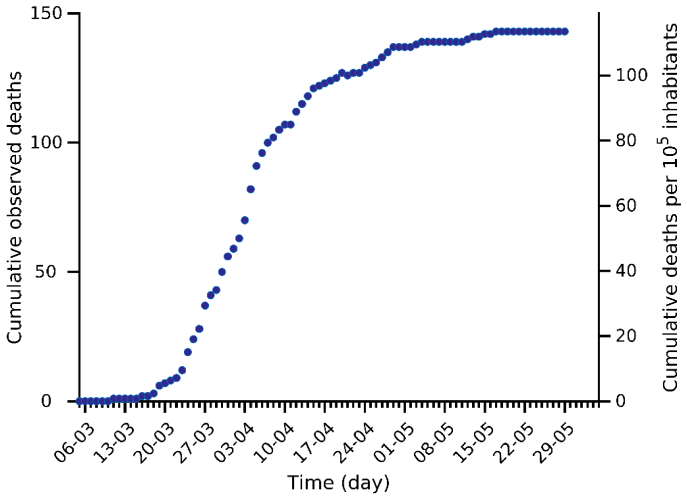


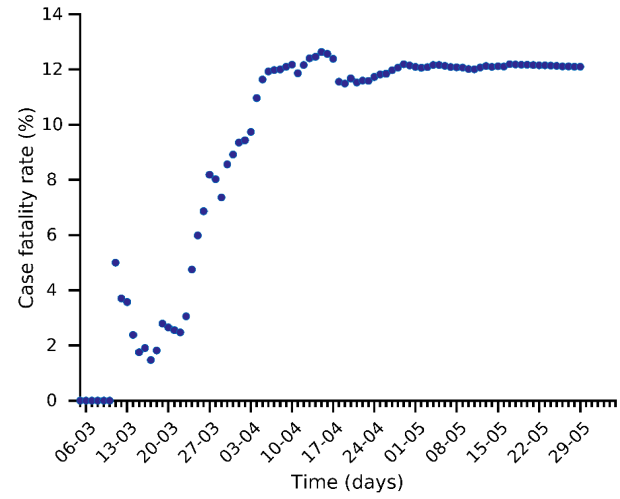




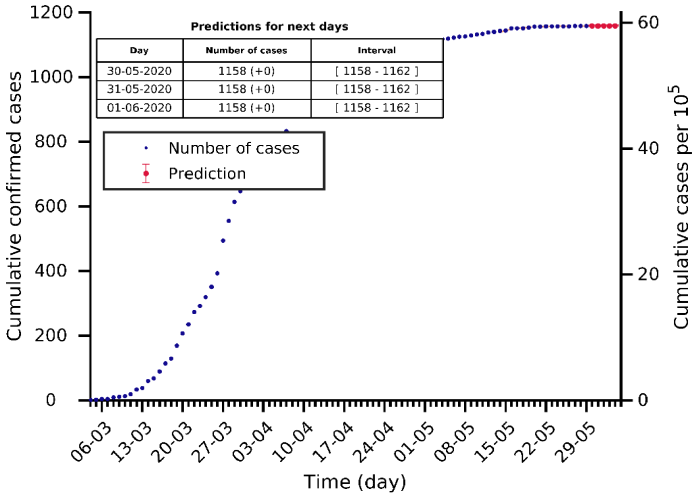


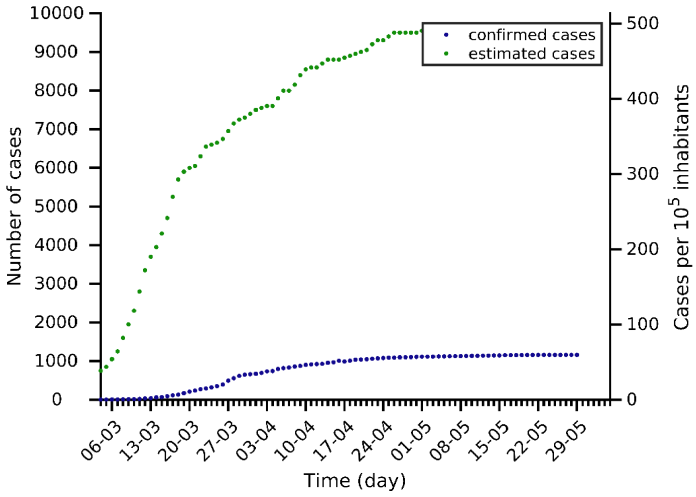


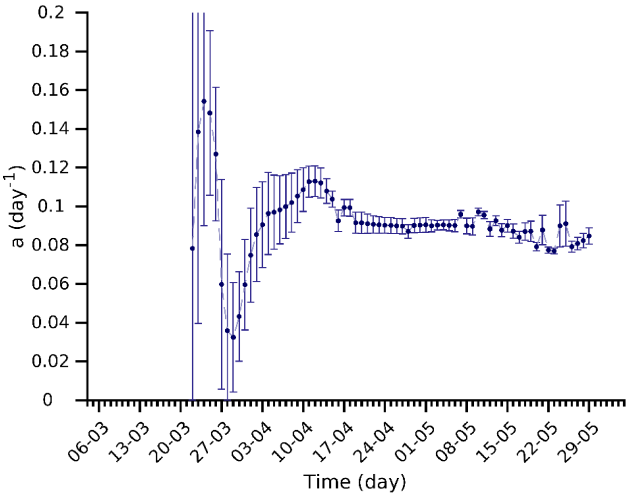


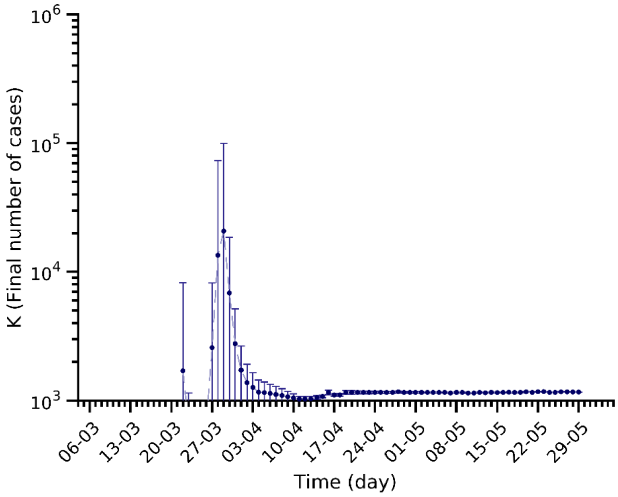


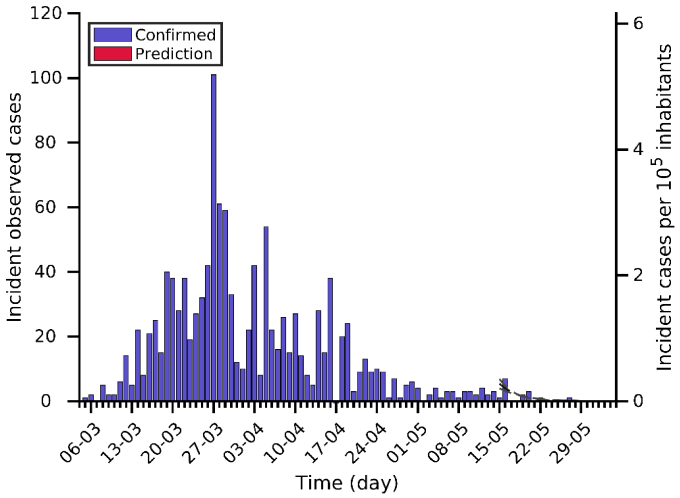
Calabria 29-05-2020. Population: 1.9M. Current cumulated incidence: 59/10^5

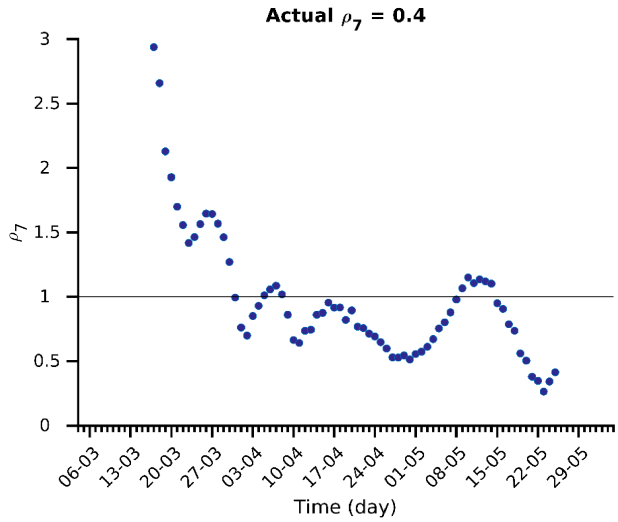


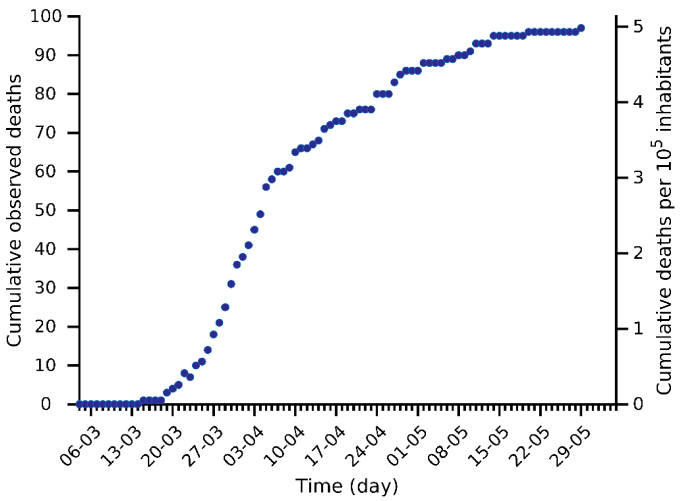


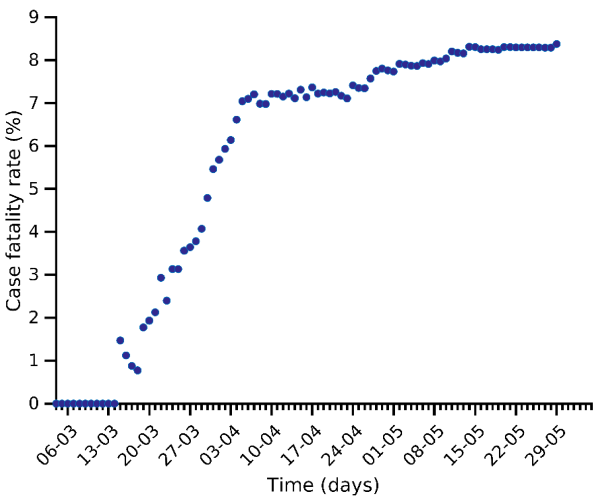




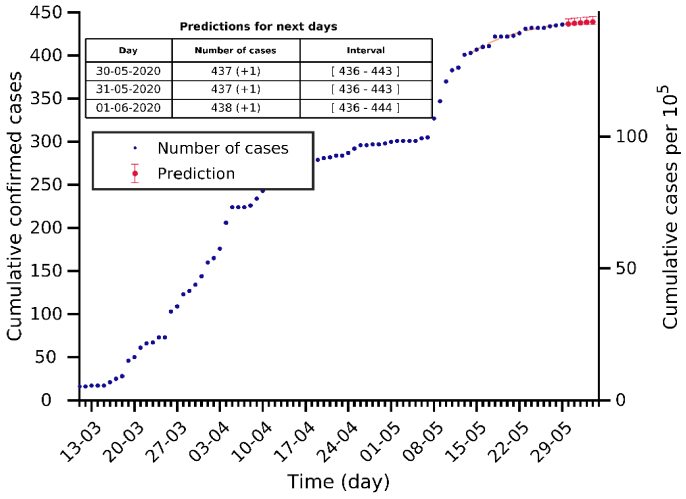


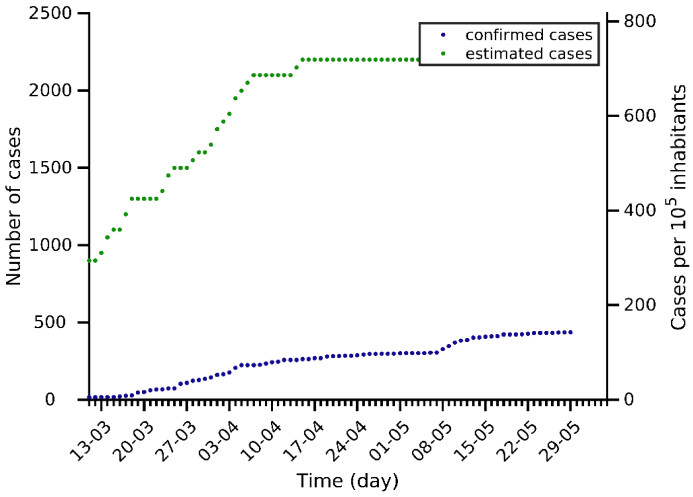


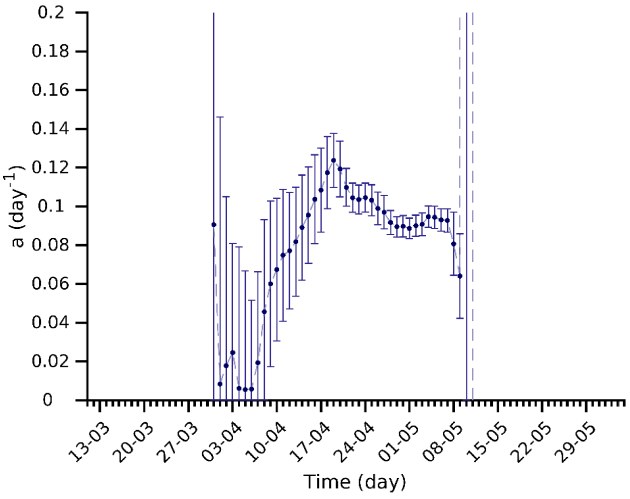


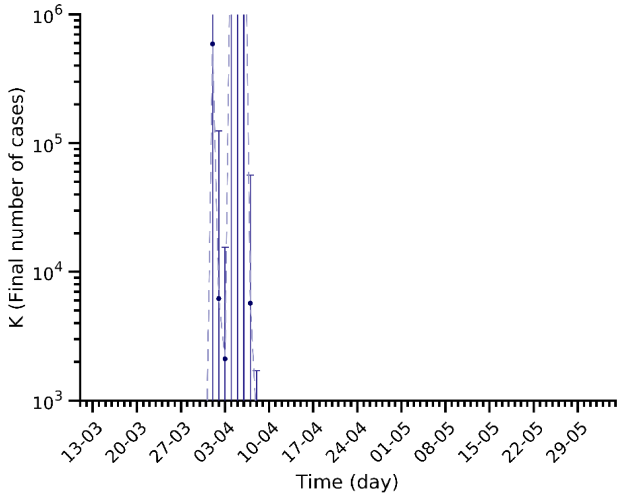


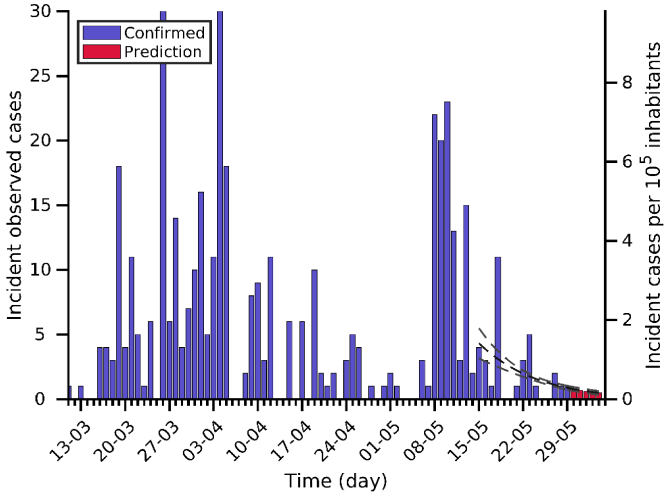
Molise 29-05-2020. Population: 0.3M. Current cumulated incidence: 142/10⁵

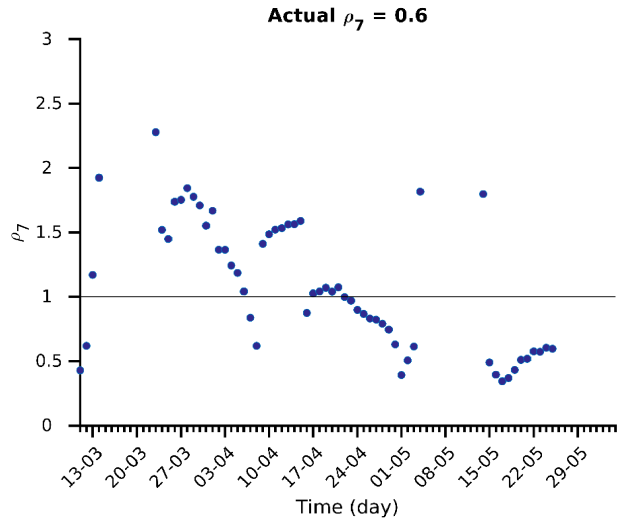


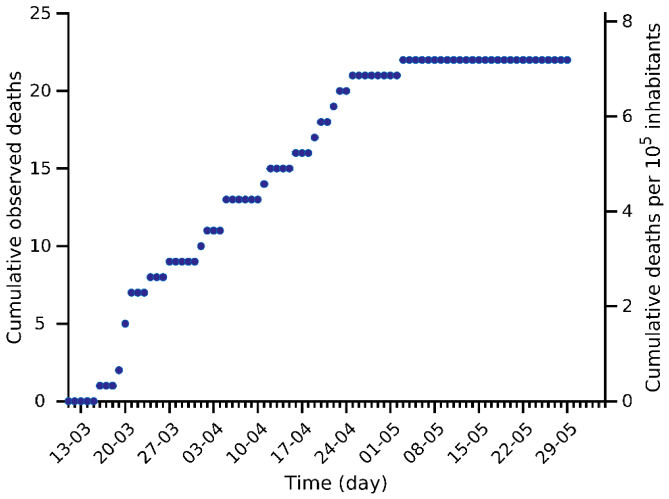


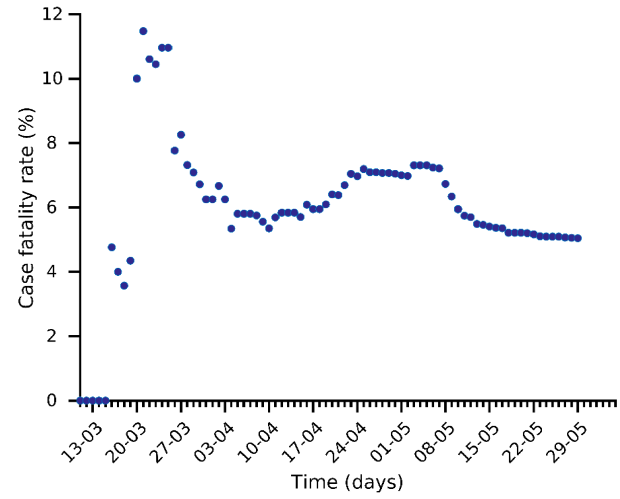




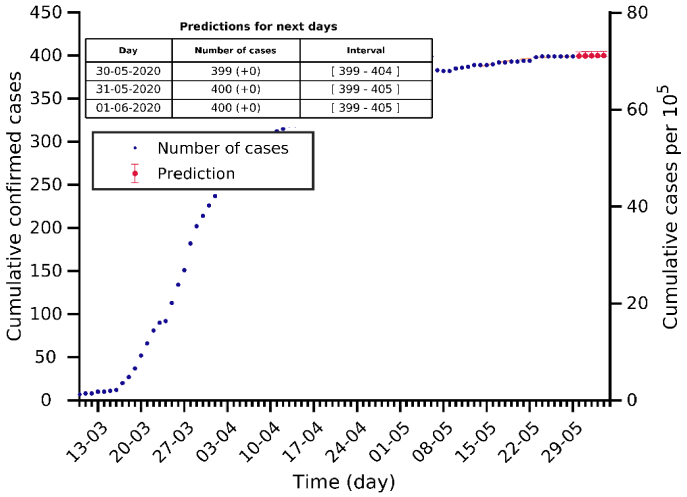


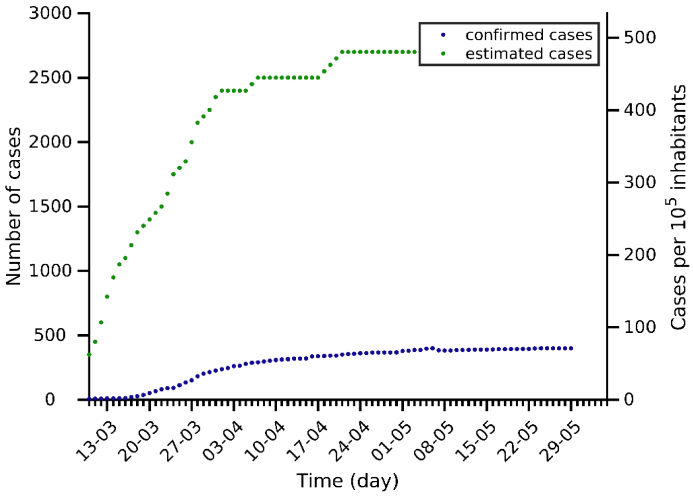


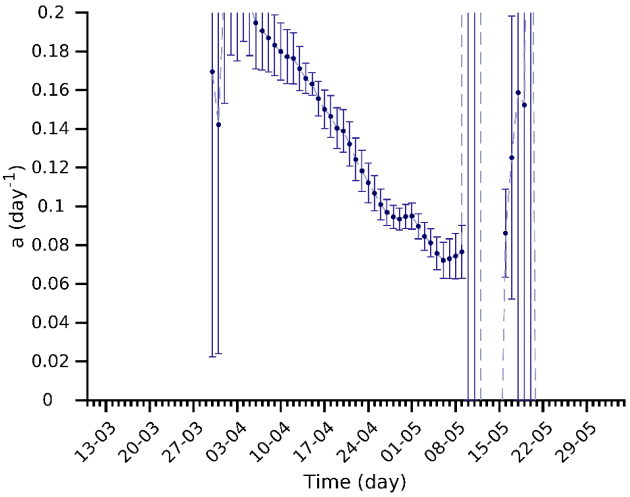


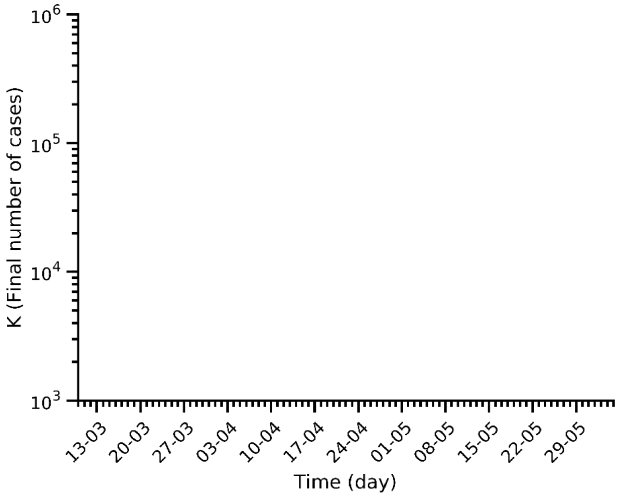


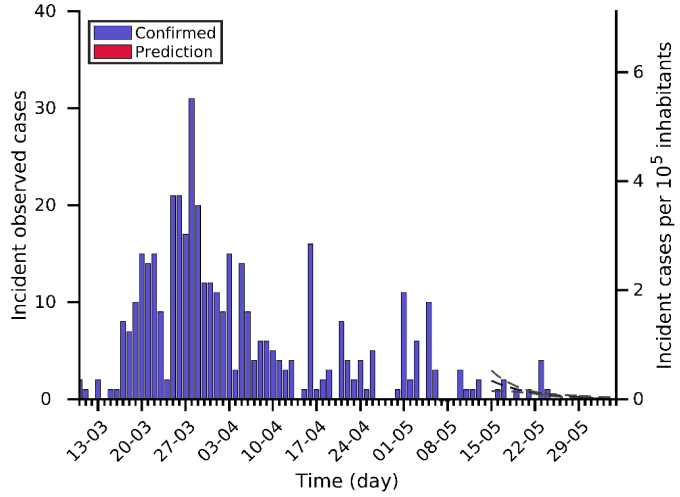
Basilicata 29-05-2020. Population: 0.6M. Current cumulated incidence: 71/10^5

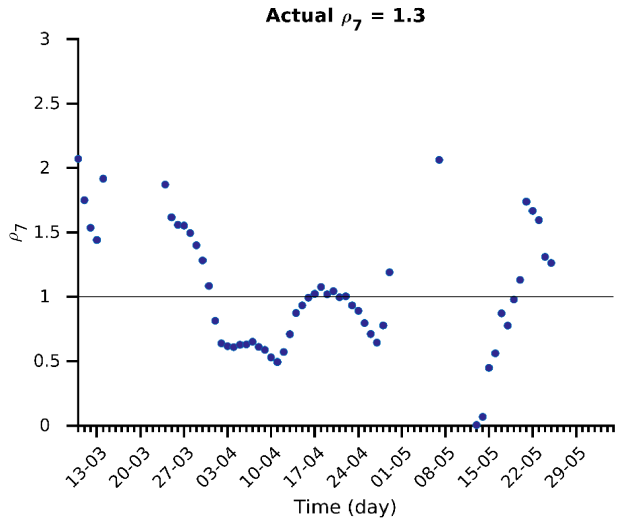


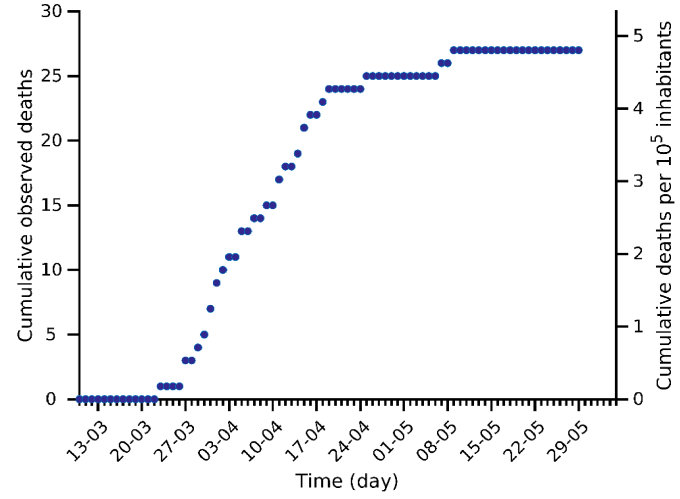


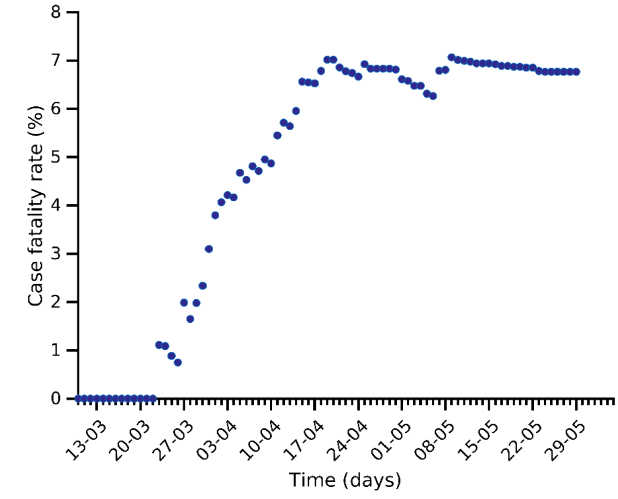












Methods

Methods

(1) Data source

Data are daily obtained from World Health Organization (WHO) surveillance reports[4], from European Centre for Disease Prevention and Control (ECDC)[5] and from Ministerio de Sanidad[6]. These reports are converted into text files that can be processed for subsequent analysis. Daily data comprise, among others: total confirmed cases, total confirmed new cases, total deaths, total new deaths. It must be considered that the report is always providing data from previous day. In the document we use the date at which the datapoint is assumed to belong, i.e., report from 15/03/2020 is giving data from 14/03/2020, the latter being used in the subsequent analysis.

(2) Data processing and plotting

Data are initially processed with Matlab in order to update timeseries, i.e., last datapoints are added to historical sequences. These timeseries are plotted for EU individual countries and for the UE as a whole:

- ✓ Number of cumulated confirmed cases, in blue dots
- ✓ Number of reported new cases
- ✓ Number of cumulated deaths

Then, two indicators are calculated and plotted, too:

- ✓ Number of cumulated deaths divided by the number of cumulated confirmed cases, and reported as a percentage; it is an indirect indicator of the diagnostic level.
- ✓ ρ: this variable is related with the reproduction number, i.e., with the number of new infections caused by a single case. It is evaluated as follows for the day before last report (*t-1*):

$$\rho(t-1) = \frac{N_{new}(t) + N_{new}(t-1) + N_{new}(t-2)}{N_{new}(t-5) + N_{new}(t-6) + N_{new}(t-7)}$$

where $N_{new}(t)$ is the number of new confirmed cases at day t.

(3) Classification of countries according to their status in the epidemic cycle

The evolution of confirmed cases shows a biphasic behaviour:

(I) an initial period where most of the cases are imported;
(II) a subsequent period where most of new cases occur because of local transmission.

Once in the stage II, mathematical models can be used to track evolutions and predict tendencies. Focusing on countries that are on stage II, we classify them in three groups:

- Group A: countries that have reported more than 100 cumulated cases for 10 consecutive days or more;
- Group B: countries that have reported more than 100 cumulated cases for 7 to 9 consecutive days;
- Group C: countries that have reported more than 100 cumulated cases for 4 to 6 days.

[4] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports
[5] https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases
[6] https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov-China/situacionActual.htm https://github.com/datadista/datasets/tree/master/COVID%2019 , https://covid19.isciii.es/

(4) Fitting a mathematical model to data

Previous studies have shown that Gompertz model[7] correctly describes the Covid-19 epidemic in all analysed countries. It is an empirical model that starts with an exponential growth but that gradually decreases its specific growth rate. Therefore, it is adequate for describing an epidemic that is characterized by an initial exponential growth but a progressive decrease in spreading velocity provided that appropriate control measures are applied.

Gompertz model is described by the equation:

$$N(t) = K\, e^{-ln\left(\frac{K}{N_0}\right)\cdot e^{-a\cdot(t-t_0)}}$$

where *N(t)* is the cumulated number of confirmed cases at *t* (in days), and N_0 is the number of cumulated cases the day at day t_0. The model has two parameters:

- ✓ *a* is the velocity at which specific spreading rate is slowing down;
- ✓ *K* is the expected final number of cumulated cases at the end of the epidemic.

This model is fitted to reported cumulated cases of the UE and of countries in stage II that accomplish two criteria: 4 or more consecutive days with more than 100 cumulated cases, and at least one datapoint over 200 cases. Day t_0 is chosen as that one at which *N(t)* overpasses 100 cases. If more than 15 datapoints that accomplish the stated criteria are available, only the last 15 points are used. The fitting is done using Matlab's Curve Fitting package with Nonlinear Least Squares method, which also provides confidence intervals of fitted parameters (*a* and *K*) and the R^2 of the fitting. At the initial stages the dynamics is exponential and *K* cannot be correctly evaluated. In fact, at this stage the most relevant parameter is *a*. Fitted curves are incorporated to plots of cumulative reported cases with a dashed line. Once a new fitting is done, two plots are added to the country report:

- ✓ Evolution of fitted *a* with its error bars, i.e., values obtained on the fitting each day that the analysis has been carried out;
- ✓ Evolution of fitted *K* with its error bars, i.e., values obtained on the fitting each day that the analysis has been carried out; if lower error bar indicates a value that is lower than current number of cases, the error bar is truncated.

These plots illustrate the increase in fittings' confidence, as fitted values progressively stabilize around a certain value and error bars get smaller when the number of datapoints increases. In fact, in the case of countries, they are discarded and set as "Not enough data" if $a>0.2\ day^{-1}$, if $K>10^6$ or if the error in K overpasses 10^6.

It is worth to mention that the simplicity of this model and the lack of previous assumptions about the Covid-19 behaviour make it appropriate for universal use, i.e., it can be fitted to any country independently of its socioeconomic context and control strategy. Then, the model is capable of quantifying the observed dynamics in an objective and standard manner and predicting short-term tendencies.

(5) Using the model for predicting short-term tendencies

The model is finally used for a short-term prediction of the evolution of the cumulated number of cases. The predictions increase their reliability with the number of datapoints used in the fitting. Therefore, we consider three levels of prediction, depending on the country:

[7] Madden LV. Quantification of disease progression. Protection Ecology 1980; **2**: 159-176.

- Group A: prediction of expected cumulated cases for the following 3-5 days[8];
- Group B: prediction of expected cumulated cases for the following 2 days;
- Group C: prediction of expected cumulated cases for the following day.

The confidence interval of predictions is assessed with the Matlab function `predint`, with a 99% confidence level. These predictions are shown in the plots as red dots with corresponding error bars, and also gathered in the attached table. For series longer than 9 timepoints, last 3 points are weighted in the fitting so that changes in tendencies are well captured by the model.

(6) Estimating non-diagnosed cases

Lethality of Covid-19 has been estimated at around 1 % for Republic of Korea and the Diamond Princess cruise. Besides, median duration of viral shedding after Covid-19 onset has been estimated at 18.5 days for non-survivors[9] in a retrospective study in Wuhan. These data allow for an estimation of total number of cases, considering that the number of deaths at certain moment should be about 1 % of total cases 18.5 days before. This is valid for estimating cases of countries at stage II, since in stage I the deaths would be mostly due to the incidence at the country from which they were imported. We establish a threshold of 50 reported cases before starting this estimation.

Reported deaths are passed through a moving average filter of 5 points in order to smooth tendencies. Then, the corresponding number of cases is found assuming the 1 % lethality. Finally, these cases are distributed between 18 and 19 days before each one.

[8] At this moment we are testing predictions at 4 days for countries with more than 100 cumulated cases for 13-15 consecutive days, and 5 days for 16 or more days.

[9] Zhou et al., 2020. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet; March 9, doi: 10.1016/S0140-6736(20)30566-3